Characterising two newly identified Arabidopsis thaliana SUMO proteases

CROLEY, FENELLA,MIKAE

Characterising two newly identified *Arabidopsis thaliana* SUMO proteases

Fenella Mikae Croley

Department of Biosciences
Durham University
Supervised by Professor Ari Sadanandom
March 2019

Thesis Submitted for The Degree of Masters By Research In Biological Sciences

Abstract

As food production efforts are under escalating threat particularly with abiotic and biotic stresses depleting crop yield, there is an increasing need to understand and manipulate the plant stress signalling pathways to generate stress-resilient crops. Recently, the post-translational modification (PTM) system, Small Ubiquitin-like Modifier (SUMO), has been shown to regulate a wide spectrum of plant adaptation processes. The research in this thesis explores our current knowledge of the SUMO pathway and investigates the SUMO proteases regulating deSUMOylation. Two proteases from a newly discovered class of SUMO proteases, deSUMOylating Isopeptidases (DeSis), in *Arabidopsis thaliana* (Arabidopsis) were extensively investigated in this study. The *At4g25660* (AT60) and *At4g25680* (AT80) DeSi proteases, displayed similar characteristics to one another and were both found to localise outside the nucleus towards the plasma membrane. An *in vitro* deSUMOylation assay displayed signs of the SUMO protease activity of the AT60 protease. Although functional redundancy was speculated between the two DeSi proteases, findings suggested unequal redundancy was more likely with AT80 being more important. Double knockout (KO) AT60-AT80 mutants using the clustered regularly interspaced short palindromic repeats (CRISPR) system and single AT60 and AT80 overexpressing transgenics were generated and subjected to stress-response assays. AT60-AT80KO mutants were hypersensitive to the presence of the stress modulator phytohormone, abscisic acid (ABA), and the pathogen response elicitor, flg22. Overexpressing lines displayed either no difference or increased tolerance to the stress elicitors relative to wild-type (WT) plants. The findings provided evidence that the AT60 DeSi protease was implicated in negatively regulating ABA signalling and plant immune responses. The AT80 protease was found to play a regulatory role in ABA and immune signalling responses, as well as showing potential implications in pathogen-induced guard cell responses. This study provides evidence the two DeSi proteases play a significant role in regulating the stress-induced growth and defence responses in Arabidopsis.

Acknowledgements

I would like to thank my supervisor Dr. Sadanandom for all the help, guidance and support I received over my Masters' year. I also owe a great deal to Linda Millyard, Rebecca Morrell, Mahsa Movahedi and Beatriz Orosa for teaching me all the essential technical skills required to perform all the protocols and experiments conducted in this study. In addition, a special thanks to Rebecca Morrell and Catherine Gough for helping me with follow-up experiments and data collection. Thank you to all the lab personnel who assisted me throughout the year and for creating a welcoming environment to work in. Lastly, I would like to thank Linda Millyard and Harry Perry for reading through the thesis document and providing valuable feedback.

Table of Contents

List of Figures

List of Tables

Abbreviation List

ABA	Abscisic Acid
APS	Ammonium Persulphate
BP	Base pairs
CaMV	Cauliflower mosaic virus
cDNA	copy deoxyribonucleic acid
Col-0 WT	Columbia-0 wild type
DeSi	deSUMOylation isopeptidase
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
dNTP	deoxynucleotide triphosphate
DTT	dithiothreitol
E.coli	Escherichia coli
ECL	enhanced chemoluminescence
EDTA	ethylenediaminetetraacetic acid
ELS1	ESD4-LIKE SUMO PROTEASE 1
ESD4	EARLY IN SHORT DAYS 4
FLG22	flagellin peptide fragment of 22 amino acids
FLS2	FLAGELLIN SENSING 2
FPLC	fast protein liquid chromatograph
GA	gibberellin
GFP	Green fluorescent protein
GST	glutathione-S-transferase
HA	hemagglutinin
His	histidine
IPTG	isopropylthiogalactoside
kDa	kilodaltons
LB	Luria-Bertani
MMS21	METHYL METHANE SULFONATE SENSITIVITY 21
MS	Murashige and Skoog
NLS	Nuclear localisation signal
OD	optical density
OTS1/2	OVERLY TOLERANT TO SALT 1/2
PAMP	pathogen-associated molecular pattern
PBS	phosphate-buffered saline
PCR	polymerase chain reaction
PMSF	phenylmethylsulfonyl fluoride
PTI	pathogen-triggered immunity
PTM	post translational modification
qPCR	quantitative polymerase chain reaction

RNA	ribonucleic acid
RT-PCR	reverse transcription polymerase chain reaction
SA	salicylic acid
SAE1/2	SUMO ACTIVATING ENZYME 1/2
SCE1	SUMO CONJUGATING ENZYME 1
SDS	sodium dodecyl sulphate
SDS-PAGE	sodium dodecyl sulphate polyacrylamide gel electrophoresis
SENP	SENTRIN-SPECIFIC PROTEASE
SIM	SUMO interaction motif
SIZ1	SAR AND MIZ 1
SOC	super optimal broth with catabolite repression
SPF1/2	SPF1/2 SUMO PROTEASE RELATED TO FERTILITY 1/2
SUMO	SMALL UBIQUITIN-LIKE MODIFIER
TAE	tris acetic acid EDTA
TAIR	The Arabidopsis Information Resource
TBST	tris buffered saline tween
TEMED	tetramethylethylenediamine
UBL	ubiquitin-like modifier
ULP	UBIQUITIN-LIKE PROTEASE
UV	ultra violet
XopD	*Xanthomonas* outer protein D
YFP	yellow fluorescent protein

1. Introduction

1.1. Introduction to Post-Translational Modifications (PTMs)

As sessile organisms, plants must adjust their physiology and development to survive being subjected to an array of environmental stresses. To achieve these rapid adaptive responses and mitigate potential damage from external stimuli, plants have evolved key signalling systems to permit the intricate regulation of multiple cellular pathways and biological events. PTM of proteins are one of these vital mechanisms which expand proteome diversity, complexity and functionality, allowing for rapid adaptive responses in plants without altering protein synthesis or turnover rates (Kwon *et al.*, 2006; Hashiguchi and Komatsu, 2016). These modifications regulate protein functionality, localisation, stability and dynamic interactions with other proteins and molecules, impacting signalling pathways and gene expression (Friso and Wijk, 2015). PTMs are covalent processes which alter the primary structure of proteins either permanently, such as with the proteolytic cleavage of a signal peptide, or as a reversible addition and removal of a functional group, like with phosphorylation (Kwon *et al.*, 2006; Webster and Thomas, 2012). As PTMs are critical in mediating the interaction between the plant and its environment, they have become of high interest within the area of plant stress biology.

Ubiquitination is one PTM system that has been extensively studied owing to its crucial role implicated in various aspects of plant biology including growth, development and responses to environmental stimuli. In Arabidopsis, approximately 5% of the proteome encodes components related to the ubiquitin modification system (Smalle and Vierstra, 2004). The ubiquitination pathway involves the covalent conjugation of ubiquitin by its C-terminal carboxyl group to the target protein's lysine residue, thereby modifying the protein's stability, localisation or function (Bartel and Citovsky, 2012). Ubiquitin, a small protein composed of 76 amino acids, attaches to its target substrate through an adenosine triphosphate (ATP)-dependent conjugation cascade driven by three enzymes (Sharma *et al*., 2016). The ubiquitination process begins with the ubiquitin activating enzyme (E1) catalysing the adenylation of the C-terminal carboxyl group of ubiquitin in an ATP-dependent manner. The activated ubiquitin is transferred to the ubiquitin conjugating enzyme (E2), where it is subsequently conjugated to the target substrates lysine residue either directly or mediated by the ubiquitin protein ligase (E3). The E3 enzyme, which is encoded by approximately 90% of the ubiquitin proteome in Arabidopsis, is responsible for conferring target substrate specificity in either situation (Moon, Parry and Estelle, 2004). Deubiquitinating enzymes (DUBs), which are cysteine proteases, are responsible for removing ubiquitin from tagged substrates permitting ubiquitin to be recycled and used for subsequent rounds of ubiquitination (Sharma *et al.*, 2016).

Reiteration of the ubiquitination process can result in polyubiquitination, which involves the contribution of a fourth ubiquitin enzyme; the ubiquitin elongating enzyme (E4). Additional ubiquitin molecules can be attached to one of the seven lysine sites situated on the preceding ubiquitin acting as a substrate itself, therefore creating multi-ubiquitin chains comprised of specific lysine-linkages between the ubiquitin moieties (Li and Ye, 2008). The degree of ubiquitination and the selected lysine residue for chain formation determines the fate of the target substrate. For instance, 26S proteasomal degradation, the most widely recognised function of the ubiquitination pathway, occurs when the substrate is poly-ubiquitinated via lysine 48 residue chains. Whereas, multi-ubiquitin

chains formed by lyinse-63 linkages have been shown to play a role in DNA replication, repair, protein synthesis and iron deficiency impacting root growth (Bartel and Citovsky, 2012; Sharma *et al.*, 2016).

1.2. Introduction to the Small Ubiquitin-Related Modifier (SUMO) System

1.2.1. Overview of SUMO

With extensive studies demonstrating the importance of the ubiquitination pathway in plants, there is growing interest in identifying and elucidating other proteinaceous PTMs. One of the modification systems which has received significant attention from the plant community is the SUMO system. With resemblance to the ubiquitination pathway, the SUMO protein, an important member of the ubiquitin-like modifiers (UBLs) superfamily, covalently attaches to and detaches from particular lysine residues on target proteins altering the fate of the protein (Mishra *et al.*, 2017). The major difference between the two PTM systems is that SUMOylation does not target proteins for proteasomal degradation and is instead found to be implicated in a diverse scope of biological pathways. Recent studies have confirmed SUMO as one of the master regulators of plant adaptation processes demonstrating its pivotal role in regulating plant growth, developmental and defence responses (Verma *et al*., 2017).

1.2.2. SUMO Protein Structure

The SUMO protein is much more structurally complex in comparison to other post-translational modifiers such as methyl or acetyl groups (Gill, 2004). SUMO proteins (12 kDa), approximately 100 amino acids in length, have a compact core sequence and varying N-terminal and C-terminal sequences. The 3-dimensional structure of SUMO resembles that of ubiquitin (9 kDa) as can be observed in Figure 1.1, despite having less than 18% similarity in their amino acid sequence (Park and Seo, 2007). Both SUMO and ubiquitin proteins possess the characteristic motif known as the ubiquitin fold, which follows a β-β-α-β-β-α-β strand arrangement as the α helix, surrounded by 5 β-strands, crosses over the molecule diagonally (Park *et al.*, 2011). Furthermore, both proteins possess a di-glycine (GG) motif at the C-terminus; a crucial feature required for conjugation to the target substrate (Dohmen, 2004). More importantly, however, are the dissimilarities between SUMO and ubiquitin, which includes the significant variation in surface charge distributions, as well as SUMO proteins possessing a protruding N-terminal extension that is absent in the ubiquitin protein. These distinctive features of SUMO most likely account for the protein's unique functions and association with specific enzymes involved in the conjugation and de-conjugation process (Gill, 2004).

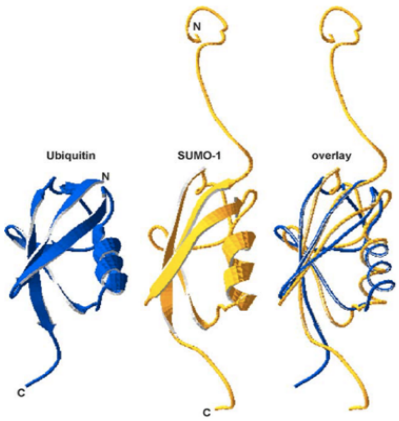


Figure 1.1 – Structural comparison of ubiquitin and *Homo sapiens* (human) SUMO-1. Both proteins share a characteristic β-grasp fold comprising of the compact core and a C-terminal GG motif; the crucial feature to their conjugation properties. The long and flexible N-terminal extension is unique to SUMO-1. The structure of ubiquitin and SUMO-1 was defined using X-ray crystallography (Vijay-Kumar *et al.*, 1987) and by nuclear magnetic resonance spectroscopy (NMR). Adapted from Dohmen (2004).

1.2.3. SUMO Isoforms

In the Arabidopsis genome, eight SUMO genes have been discovered (AtSUMO1 – AtSUMO8). However, only four of these genes, AtSUMO1, 2, 3 and 5, encode functionally divergent variations of the SUMO protein. AtSUMO1 shares the highest protein sequence identity with AtSUMO2 (89%), whereas with the other functional paralogs, AtSUMO3 (48%) and AtSUMO5 (35%), are much more distantly related (Verma *et al.*, 2018). Despite varying in sequence similarity, all members of the Arabidopsis SUMO family share several highly conserved residues including the C-terminal glycine where target substrate conjugation takes place. However, these isoforms diverge significantly from one another displaying heterogeneous biochemical and functional properties. AtSUMO1 and AtSUMO2 (AtSUMO1/2) and their corresponding conjugates have been detected in both cytoplasmic and nuclear compartments, whereas AtSUMO3 and AtSUMO5 (AtSUMO3/5) proteins have been found to only concentrate in specific tissues such as leaf vasculature and roots (Park *et al.*, 2011). Furthermore, preferential conjugation of AtSUMO1/2 over AtSUMO3/5 has been illustrated when comparing *in vitro* conjugation rates between the isoforms. This divergence in conjugation rates has been ascribed to the absence of conserved residues in the AtSUMO3/5 proteins. These residues are involved in the initial step of the SUMO-conjugation cascade, subsequently impacting thioester-bond formation (Castaño-Miquel *et al.*, 2011). On the other hand, studies have proven the participation of AtSUMO1, 2 and 3 in SUMO conjugation *in vivo*. However, as the expression and abundance of AtSUMO1/2 are significantly greater than AtSUMO3, they are the best studied and therefore the most canonical isoforms of the AtSUMO protein family (Novatchkova *et al.*, 2012). Studies have confirmed AtSUMO1/2 conjugates rapidly increase in response to environmental stresses, in particular, heat shock and H2O2 treatment (Kurepa *et al.*, 2003). The concomitant inactivation, double KO, of AtSUMO1/2 was embryo lethal, thereby implying the two isoforms could be functionally redundant. This research underlined the significance of both AtSUMO1/2 in the regulation of plant growth, development, and stress responses.

1.2.4. SUMO Pathway and Machinery

The SUMO pathway is comprised of three major phases depicted in Figure 1.2. This pathway begins with SUMO maturation, followed by SUMOylation, a stepwise enzymatic cascade conjugating the SUMO moiety to the target protein, and lastly deSUMOylation, which removes SUMO from the substrate for further conjugation cycles.

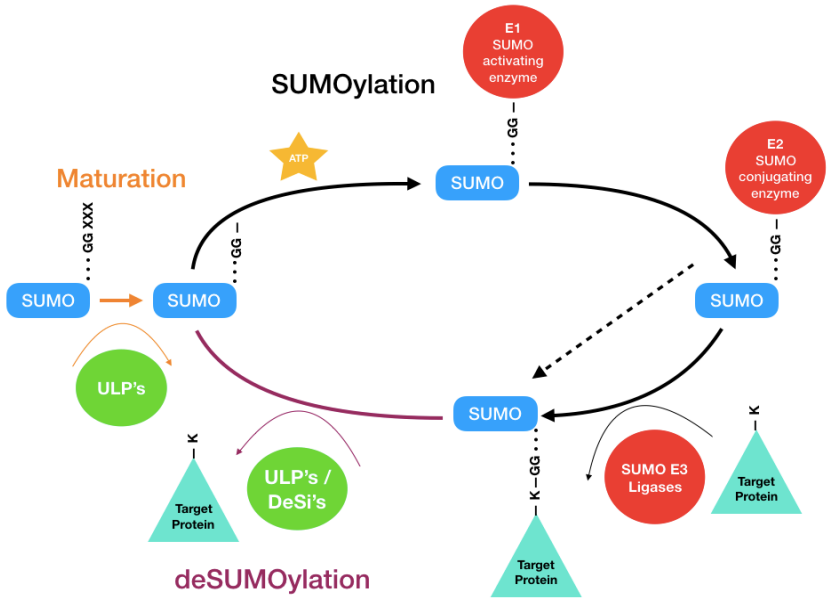


Figure 1.2 – SUMO pathway and machinery. The pathway begins with the precursor SUMO protein cleaved by ubiquitin-like proteases (ULPs) producing mature SUMO. Subsequently, SUMO is activated by E1 using ATP, then conjugated by E2 and finally ligated to the target substrate via the E3 enzyme. SUMO conjugates to the lysine residue of the target substrate via E3 or even in some cases, directly via the E2 enzyme. DeSUMOylation by ULP or DeSi SUMO proteases, remove SUMO from the target substrate to be recycled for subsequent rounds of SUMO conjugation.

Similar to ubiquitin, SUMO is synthesised as a longer precursor protein that must be processed to its mature form for post-translational conjugation. ULPs are a class of SUMO-specific cysteine proteases in plants that utilise their endopeptidase activity to facilitate the maturation of the SUMO precursor protein. These proteases recognise the C-terminal GG motif in SUMO proteins and cleaves approximately 10 amino acids directly after the GG motif, thereby exposing the reactive carboxyl group of the second glycine residue (Park and Seo, 2007; Elrouby, 2015). Members of the ULP family are also involved in SUMO de-conjugation, where the SUMO protease cleaves the isopeptidase linkage between the SUMO moiety and the target substrate. To date, eight ULPs have been discovered in Arabidopsis, however, there is little studied regarding these proteases (Park *et al.*, 2011).

Once SUMO has been processed to its mature form, the modifier conjugates to the target substrate via a step-wise enzymatic cascade resembling that of the ubiquitination pathway. This process, known as SUMOylation (depicted in Figure 1.2), is catalysed by three specific enzymes: SUMO activating enzyme (E1), SUMO-conjugating enzyme (E2), and SUMO ligase (E3). The first step of SUMOylation is catalysed by the SUMO E1 enzyme. In Arabidopsis, this enzyme is heterodimeric comprising of a smaller 40kDa subunit (SAE1) and a larger 70kDa protein (SAE2)(Park and Seo, 2007). To initiate SUMOylation, the SUMO protein is first activated. The C-terminus carboxyl group within the GG motif of the SUMO protein reacts with ATP forming adenylated SUMO. The thiol group in the catalytic cysteine residue of SAE2 then interacts with the SUMO adenylate producing a high-energy thioester bond between the SUMO E1 enzyme and the carboxyl group of SUMO, subsequently freeing adenosine monophosphate (AMP) (Park *et al.*, 2011). Whilst the larger SAE2 subunit is encoded by a single gene in Arabidopsis, SAE1 is represented by two functional genes; SAE1a and SAE1b. Both isoforms are competent for SUMO activation and were thought to be functionally redundant. However recent studies demonstrated higher conjugation efficiencies in the SAE1a variant as opposed to SAE1b, as well as establishing SAE1a is required for maintaining homeostasis during SUMO conjugation. These results, therefore, suggest the SAE1 subunit may be implicated in the downstream regulation of SUMOylation (Budhiraja *et al.*, 2009; Castaño-Miquel *et al.*, 2013).

With SUMO activated, the moiety is then transferred to the active site cysteine residue on the SUMO E2 enzyme in a transesterification reaction. This results in a SUMO-E2 thioester intermediate. In Arabidopsis, a single E2 gene has been identified; SCE1 (Elrouby, 2015). The E2 enzyme can directly conjugate SUMO to target proteins by mediating the formation of an isopeptide linkage between the carboxyl-terminal glycine of SUMO and the ε-amine group of a lysine residue within the substrate. This occurs under circumstances where the target substrate contains a specific consensus motif known as the SUMOylation consensus sequence: ΨKxE (Ψ = hydrophobic amino acids; K = SUMO target lysine; x = any amino acid; E = acidic amino acids). This is a distinguishing characteristic from the ubiquitination system which relies only on E3 ligases for substrate

specificity. Instead, substrate specificity is conferred by the E2 enzymes ability to recognise the SUMOylation consensus motif exposed on the surface of the target substrate (Colby, 2006; Mazur and van den Burg, 2012).

Although studies *in vitro* were able to prove that E1 and E2 enzymes are sufficient for the SUMOylation of different substrates (Bernier-Villamor *et al*., 2002) it is believed SUMO E3 ligases are necessary in facilitating the process *in vivo*. E3 ligases can act as a scaffold bringing the SUMO-charged E2 enzyme and target substrate in close proximity by providing binding interfaces for both. Alternatively, it has been suggested certain E3 ligases can stimulate SUMO transfer to the substrate by interacting with the E2 enzyme and imposing conformational constraints (Novatchkova *et al.*, 2012). As a result, these ligases are thought to accelerate the rate of SUMO conjugation and influence the extent of SUMOylation as well as determine substrate specificity, particularly for substrates that lack the SUMOylation consensus motif (Gill, 2004). Currently, only two E3 ligases have been identified in Arabidopsis; METHYL METHANE SULFONATE SENSITIVITY PROTEIN 21 (MMS21) and SAP AND MIZ1 (SIZ1). Both enzymes possess the highly conserved E2 enzyme interaction domain identified in most SUMO ligases known as the SP-RING domain (Novatchkova *et al.*, 2012). Two additional ligase proteins encompassing the SP-RING domain have also been identified in the Arabidopsis genome; PROTEIN INHIBITOR OF ACTIVATED STAT LIKE 1 (PIAL1) and 2 (PIAL2). These ligases were found to contribute towards the formation of SUMO chains and exhibited no functional redundancy to E3 ligases. Whether a substrate is mono- or poly-SUMOylated is believed to be influenced by the innate properties of the substrate (Tomanov *et al.*, 2014).

In addition to covalent SUMO conjugation, non-covalent binding of a target substrate to SUMO can occur through a SUMO-interacting motif (SIM) present in the substrate. This binding mechanism is regarded analogous to the ubiquitin system where ubiquitin covalently conjugating to the target mobilises ubiquitin-binding proteins to bind non-covalently (Park *et al.*, 2011). SIMs are defined by a hydrophobic consensus sequence comprising clusters of valine, isoleucine and leucine residues and groups of acidic or phosphorylated residues which correspond to specific SUMO isoforms (Yates *et al*., 2016). One study was successful in proving the crucial role of these non-covalent interactions in the regulation of plant growth under high salinity. SUMOylated DELLA can interact with GA INSENSITIVE DWARF1 (GID1) through its SIM, resulting in the sequestration of GID1 by SUMO conjugated-DELLAs. As a consequence, non-SUMOylated DELLAs increase in the cell, thereby permitting advantageous growth repression during stress conditions (Conti *et al.*, 2014). It is believed these non-covalent interactions could be driving complex protein formations or alternatively impeding interactions by covering partner binding sites. However, the functionality of these non-covalent interactions and the significance of the SIM are yet to be fully elucidated and still undergoing research.

SUMO conjugation to protein substrates is a reversible system to maintain equilibrium in SUMO signalling by freeing the SUMO moiety. The process entailing the de-conjugation of SUMO from a target substrate is known as deSUMOylation; another critical phase of the SUMO pathway (Figure 1.2). DeSUMOylation is catalysed by a family of cysteine proteases termed SUMO proteases, which specifically cleave the isopeptide linkage between the SUMO monomer and substrate. ULPs are a class of SUMO proteases that dually function in both SUMO maturation and de-conjugation, utilising their endopeptidase and isopeptidase activity, respectively (Augustine and Vierstra, 2018; Garrido *et al*., 2018). As bioinformatic studies have collectively identified more SUMO proteases than SUMO E3 ligases in plants, there is a strong indication that SUMO proteases regulate

the specificity of target substrates (Grau-Bové *et al.*, 2015; Orosa *et al.*, 2018). The topic of deSUMOylation with regard to recent studies identifying new SUMO protease classes will be further elaborated upon in section 1.4.

1.3. Role of SUMO in Plants

Analyses of mutants through gain-of-function and loss-of-function have provided emerging insight into the pivotal role SUMOylation plays in regulating plant development, signalling and adaptive responses. There is strong evidence of SUMOylation involvement in plant development as findings have shown KO mutations in either SAE1/2, SCE1, SIZ1 and MMS21, or SUMO1/2 as embryonic lethal (Saracco *et al.*, 2007; Ishida *et al.*, 2012). Significant focus has also been placed on elucidating the relationship between the SUMO pathway and plant responses to abiotic and biotic stress, particularly as SUMO conjugation rapidly increases as result of introducing stress conditions (Verma *et al*., 2017). Studies have demonstrated SUMO conjugation and deconjugation of transcriptional complexes facilitates the accurate regulation of determinant gene expression, thereby modulating these diverse biological processes. With the diverse scope of target proteins implicated in these various biological processes including growth, flowering, and responses to external stimuli, there is a clear indication the SUMO pathway is involved in governing the growth-defence equilibrium.

SUMOylation of a substrate influences the fate of the protein. The target protein's stability, interaction dynamics, subcellular localisation and activity can be induced by SUMO conjugation. These modifications subsequently implicate the regulation of various processes including DNA repair, chromatin modification or remodelling, nuclear transport, epigenetics, and transcriptional activation and repression (Colby *et al.*, 2006; Augustine and Vierstra, 2018). For instance, a recent study demonstrated how the mono-SUMOylation of a SU(VAR)-3-9-related protein family member, SUVR2, enables its interaction with chromatin remodelling proteins resulting in transcriptional gene silencing (Luo *et al.*, 2018). Alternatively, SUMOylated proteins may undergo regulated proteolysis. Poly-SUMO chains on a substrate can function as binding domains for a class of ubiquitin E3 ligases termed SUMO-targeted ubiquitin ligases (STUbLs). To date, the Arabidopsis genome expresses six STUbLs; two of which are orthologs of mammalian and yeast STUbLs. Numerous SIMs found on the STUbL enzyme permits non-covalent binding to the poly-SUMO chains inducing dimerisation of STUbL; a prerequisite for their subsequent poly-ubiquitination of SUMO and its conjugated substrate. The ubiquitin-proteasome system then targets these proteins for proteasomal degradation (Elrouby, 2015).

Paradoxically, despite the wide-ranging impact SUMOylation has on various biological process, only a small percentage of the target protein pool is SUMOylated at any time (Augustine and Vierstra, 2018). One hypothesis explaining this contradiction is that low-level SUMOylation primes a rapid change in the state of modification as a response mechanism to particular stimuli either by enhancing or extending target SUMOylation (Elrouby, 2015). Another possibility is that protein SUMOylation is confined to highly specific time frames, cell types and protein conditions, consequently resulting in a relatively lower level of SUMOylated protein in comparison to the entire protein pool (Verma *et al*., 2018). Alternatively, low-level SUMOylation may be necessary to maintain proteostasis; a concept entailing the collective cellular pathways modulating biogenesis, folding, trafficking and degradation of proteins. SUMO proteases account for this primed cellular environment by exerting tight regulation over the deconjugation of SUMO from target substrates (Elrouby, 2015).

1.3.1. Role of SUMO in Plant Development

Plants survive challenging conditions by adjusting their growth in response to fluctuating environmental stimuli. Phytohormones and their associated signalling pathways are pivotal in modulating these adaptive responses. Extensive literature substantiates SUMO as a key player in regulating these hormonal pathways.

SUMOylation is implicated in ABA signalling; a crucial phytohormone for plants under abiotic stress. The overexpression of AtSUMO1 or 2 attenuates ABA-mediated growth inhibition and joint overexpression of the isoforms amplified the expression of the ABA-responsive genes, *RD29A* and *AtPLC1*. On the other hand, decreasing AtSCE1 expression levels increased sensitivity to the growth inhibitory effect by ABA (Castaño-Miquel *et al.*, 2013). Literature supports the notion that AtSIZ1 negatively regulates ABA signalling; a pathway dependent on the basic leucine zipper (bZIP) transcription factor, *ABI5*. *siz1* mutants displayed hypersensitivity to ABA resulting in reduced seed germination rates and seedling root growth inhibition. Miura *et al*. (2010) established that the AtSIZ1-mediated SUMOylation of ABI5 at the K391 residue results in the impediment of ABA signalling during seedling growth and germination. The SUMOylation by AtSIZ1 of another transcription factor, AtMYB30, was also found to be critical in the regulation of ABA signalling (Zheng *et al*., 2012). Interestingly, a very recent study found that the stability and function of MYB30 also depends on ubiquitination. ABA induces the ubiquitin E3 ligase, RHA2b, to target MYB30 for degradation via the 26S proteasome. SUMOylation of MYB30 on the other hand, protects the transcription factor from degradation. As the K283 residue on MYB30 functions as the major site for both ubiquitination and SUMOylation, this study proposes the pathways behave antagonistically in determining the stability of MYB30 as a response to ABA (Zheng *et al.*, 2018). A study by Zhang *et al*. (2013) has demonstrated the newly-identified Arabidopsis SUMO E3 ligase, MMS21, to also be implicated in ABA responses. *mms21* mutants exhibited hypersensitivity to ABA, displaying slower water loss and improved tolerance to drought conditions. Furthermore, the ABA-induced accumulation of SUMO-protein conjugates was impeded in the *mms21* mutant. The research concluded that MMS21 is implicated in drought stress responses, presumably through regulating gene expression in an ABA-mediated pathway.

SUMO also plays a role in gibberellic acid (GA) signalling, which is commonly known to function antagonistically with ABA when regulating various developmental stages (Liu and Hou, 2018). The current GA signalling model is as follows. During stress, the accumulation of growth-inhibiting DELLA proteins is opposed by GA, which promotes degradation of these growth repressing proteins. The phytohormone binds to its receptor, GID1, stimulating the interaction of GID1 with DELLA, which is subsequently ubiquitinated for proteasomal degradation (Peng *et al.*, 1997). The Skp1, Cullins, F-box (SCF) complex E3 ubiquitin ligase, which catalyses this ubiquitination, encompasses an F-box subunit encoded by the Arabidopsis *SLY1* gene which determines the substrate specificity of this E3 ligase (Kim *et al.*, 2015). Research results confirmed *SLY1* as a positive regulator of plant growth by GA signalling through stimulating the degradation of DELLA proteins. This recent study by Kim *et al.* (2015), found AtSIZ1 positively regulating GA signalling by SUMOylating SLY1. In *siz1-2* mutants, SLY1 abundance declined whilst the Arabidopsis DELLA protein, REPRESSOR OF GA (RGA), increased. The study also revealed GA significantly increased SUMO conjugation to SLY1 and confirmed the interaction between SUMOylated SLY1 with RGA. Therefore, the study concluded that the AtSIZ1-mediated SUMOylation of SLY1 stabilises and activates SLY1 for RGA

degradation, subsequently stimulating SLY1-mediated plant growth. Research over the decade has also provided evidence supporting the role of SUMO in salicylic acid (SA) (Bailey *et al.*, 2016), auxin (Miura *et al.*, 2011), cytokinin (Zhang *et al*., 2010), jasmonic acid (JA) (Srivastava *et al.*, 2018) and brassinosteroids (Khan *et al.*, 2014) signalling.

SUMO has also been directly implicated in the development of plants including flowering control and seed development. *siz1* mutants have dwarfed phenotypes with smaller leaves and exhibited early short-day flowering, partly due to the accumulation of SA induced by the repressed expression of the SA hydrolase, NahG (Jin *et al.*, 2007; Miura *et al.*, 2007). AtSIZ1 also functions independently of SA acting up-stream of the *Flowering Locus D* (*FLD*) gene; a notable activator of flowering involved in the autonomous pathway (Miura *et al.*, 2007). AtSIZ-1-mediated SUMO conjugation of FLD represses its activity resulting in the acetylation of histone 4 in the chromatin of the floral suppressor, *Flowering Locus C* (*FLC*). As a result, this upregulates FLC expression and subsequently represses the expression of floral genes (Jin *et al.*, 2007). Similarly, transgenic plants overexpressing the AtSCE protein with a mutated active site from a cysteine residue to a serine, also displayed early flowering alongside reduced growth (Tomanov *et al.*, 2013).

Interestingly, mutants of the Arabidopsis SUMO proteases including EARLY IN SHORT DAYS 4 (ESD4) and ULP1 proteases, also displayed similar early short-day flowering phenotypes like AtSIZ1, despite the two enzyme classes having opposing roles in the SUMO pathway; AtSIZ1 facilitates SUMOylation whilst SUMO proteases mediate deSUMOylation. In *esd4* mutants, which also exhibits a dwarf phenotype, FLC expression levels diminished whilst transcript levels of flowering-time genes normally suppressed by FLC heightened, subsequently driving the transition to flowering from vegetative growth (Park *et al.*, 2011). Furthermore, the double KO of the SUMO proteases, OVERLY TOLERANT TO SALT 1 and 2 (OTS1/2), also flowered early (Conti *et al.*, 2008), thereby suggesting SUMO proteases negatively regulate the switch to flowering. Furthermore, AtSUM1/2 knockdown mutants flowered early in short days. In contrast, reduced AtSUM3 expression caused late flowering whereas overexpression of AtSUM3 resulted in early flowering (van den Burg *et al.*, 2010). Studies have therefore demonstrated both SUMOylation and deSUMOylation are implicated in pathways mediating flowering in plants.

There has also been strong evidence supporting the role of SUMOylation and deSUMOylation in plant reproduction. Genetic analyses found that double mutants of the Arabidopsis E3 SUMO ligases, *mms21-1* and *siz1-2,* resulted in lethality in embryogenesis, indicating E3 ligases are required during early plant development. *mms21* mutants displayed complications in both male and female gametophytes including chromosome distributions, meiotic abortion and defective pollen tube growth (Olaofe *et al*., 2013). *siz1* mutants also exhibited reproductive issues. Following fertilisation at the globular stage, embryonic development stopped in mutant plants resulting in abnormal growth and abortion of over 50% of the mature seeds (Park and Seo, 2007).

The recently characterised nuclear-located SUMO proteases, SUMO PROTEASE RELATED TO FERTILITY 1 (SPF1) and 2 (SPF2), are also critical to gametogenesis. Liu *et al*. (2017), provided evidence that SPF1 and SPF2 regulate male and female gamete and embryo development. Single *spf1* mutant plants exhibited delayed flowering and diminished self-fertilisation as a result of abnormalities in floral structures including morphological separation of the anthers and stigma (Olaofe *et al*., 2013). Whereas, single *spf2* mutants

displayed similar phenotypes to WT Arabidopsis. Interestingly, double mutants of *spf1/2* showed extreme defects in microgametogensis, megagametogenesis and embryo development implying the two genes are functionally redundant (Liu *et al*., 2017). These findings suggest the deSUMOylating proteases, SPF1/2, may function antagonistically to MMS21. In addition, similar phenotypes were recorded in double mutants of OTS1/2, as well as late germination and leaf growth defects (Olaofe *et al*., 2013; Verma *et al*., 2018). The findings support the notion that the SUMO system is tightly linked to plant reproductive development.

1.3.2. Response to Abiotic Stress

As SUMO conjugation is significantly enhanced by abiotic stresses, this strongly supports the notion that SUMOylation is involved in plant responses and protection against environmental stresses (Park and Seo, 2007). Findings on how the mutation of genes implicated in SUMO conjugation reduces stress tolerance, highlights the significance of SUMOylation in response to abiotic stresses (Karan and Subudhi, 2012).

The increase in SUMO-conjugate levels in response to heat, cold and drought stress has primarily been shown to be AtSIZ1-dependent (Castro *et al*., 2012). A study by Miura *et al*. (2011) determined in response to phosphate (Pi) starvation conditions, AtSIZ1 negatively regulates auxin patterning for root system architecture remodelling and determinant gene expression; crucial components for the acquisition of Pi. *siz1* mutants exhibited a hypersensitive response to Pi starvation resulting in the attenuation of primary root elongation and stimulation of lateral root formation in seedlings. Further analyses also found that the expression of various auxin and Pi starvation-inducible genes were upregulated in the *siz1* mutant as opposed to WT Arabidopsis. SIZ1-mediated SUMOylation of the MYB transcription factor, PHR1, acts positively on Pi-starvation induced gene expression, therefore indicating SIZ1 negatively regulates phosphate uptake (Park *et al.*, 2011). AtSIZ1 has also been found to regulate basal thermotolerance independent of SA signalling; a pathway involved in basal and acquired thermotolerance through inducing heat shock protein (HSP) expression. Furthermore, AtSIZ1 is crucial for tolerance to cold conditions. AtSIZ1-mediated SUMO conjugation of the transcription factor inducer of CBF/DREB1 expression (ICE1), is necessary for the induction of CBF/DREB1-dependent cold signalling and freezing tolerance (Miura *et al.*, 2007; Park *et al.*, 2011).

A study by Catala *et al.* (2007) first identified how drought stress promotes the accumulation of SUMO-protein conjugates and demonstrated how *siz1* mutants exhibited a substantial loss in drought tolerance. Genomic analyses revealed how the expression of 300 out of approximately 1700 Arabidopsis drought-induced genes are mediated by AtSIZ1 via a pathway independent of ABA. Encapsulating all findings from the present literature, AtSIZ1 KO mutants display heightened sensitivity to drought, heat and low-temperature stresses, whilst constitutive expression of AtSIZ1 improved salt and cold stress tolerance (Kim *et al.*, 2016). Therefore, studies have concluded AtSIZ1 regulates stress responses in Arabidopsis. In addition to AtSIZ1, *mms21* mutants exhibited improved tolerance to drought stresses, whilst overexpressing transgenics displayed a reduction to drought tolerance. Genetic analyses showed how MMS21 deficiency resulted in the heightened expression of ABA-mediated stress-responsive genes, whereas the constitutive expression of MMS21 impeded both ABA- and drought-induced stress-responsive genes. This study proved the E3 SUMO ligase, MMS21 also plays a significant role in abiotic stress response most likely via an ABA-dependent pathway (Zhang *et al*. 2013).

1.3.3. Response to Biotic Stress

The SUMO pathway is also involved in phytopathogen infections and defence. More specifically, the deSUMOylation process was found to be a prime target during pathogenesis. This was initially observed in the bacterial pathogen, *Xanthomonas campestris* pv. *vesicatoria* (*Xcv*), which injects type III effector proteins into plant host cells that function as cysteine proteases with SUMO-substrate specificity. *Xanthomonas* outer protein D (XopD), an *Xcv* virulence factor, targets the host nucleus and mimics endogenous SUMO isopeptidase activities by cleaving SUMO-conjugated proteins. AvrBsT is an avirulent factor secreted from the same phytopathogen and was found to also possess SUMO-specific protease properties (Park *et al.*, 2011). Research has substantiated that phytopathogen-injected effector proteins mimic the activities of SUMO-specific proteases to deconjugate SUMOylated substrates, potentially a critical defence regulator, subsequently disrupting host cellular processes and facilitating pathogenesis (Hotson *et al*., 2003). These findings prove the deSUMOylation process is a key target for undermining plant immunity and altering plant defence responses.

SUMO conjugation is also directly implicated in the innate defensive response of plants against pathogens. *sum1sum2* knock-down mutants displayed heightened SA levels which corresponded to the mutants increased tolerance to the pathogen, *Pseudomonas syringae* pv. *tomato* (Pst) DC3000 (van den Burg *et al.*, 2010). Similarly, *siz1* mutants also exhibited heightened resistance to PstDC3000, attributed to the accumulation of SA and constitutive expression of pathogenesis-related genes and infection response genes. This suggests SIZ1 negatively regulates plant defence responses through a SA-dependent pathway. SA signalling is critical in plant defence against pathogens as it regulates programmed cell death and induces the expression of pathogenesis-related genes (Lockhart, 2013). Interestingly, *ots1 ots2* double mutants also exhibited increased SA levels conferring heightened resistance to PstDC3000. This is due to OTS1 and OTS2 restricting SA biosynthesis by impeding the expression of ISOCHORISMATE SYNTHASE1 (ICS1) (Bailey *et al.*, 2016). A recent study also demonstrated that by preventing SUMO conjugation through disrupting the SUMO E1-E2 interaction, transgenic plants were more susceptible to infections by *Botrytis cinerea* and *Plectosphaerella cucumerina* (Castaño-Miquel *et al.*, 2017). These findings support the notion that SUMOylation and deSUMOylation play a significant role in plant-pathogen interactions.

1.4. DeSUMOylation and SUMO-specific Proteases

DeSUMOylation is a key regulatory step in the entire SUMO pathway. The literature substantiates that SUMO proteases are involved in a range of plant biological processes and provide the specificity to substrate proteins, thereby functioning analogously to the ubiquitin E3 ligase protein (Verma *et al*., 2018). Genetic analyses further support this notion as findings show the SUMO system has a greater number of genes encoding SUMO proteases in contrast to ubiquitination, which experiences the same gene number abundance with E3 ligases (Yates *et al.*, 2016). To date, seven SUMO-specific proteases part of the wider cysteine protease family, have been identified in Arabidopsis (Table 1.1). However, only a few bona fide proteases, including ESD4 and OTS1/2, have been characterised biochemically, genetically and physiologically.

Table 1.1 – Characteristics of the seven SUMO-specific proteases in Arabidopsis. This table displays all published findings on each of the seven Arabidopsis SUMO proteases, detailing their length in amino acids (AA), subcellular localisation, predominant function, which SUMO isoform the protease is known to target for SUMO processing and deconjugation, and their biological implication.

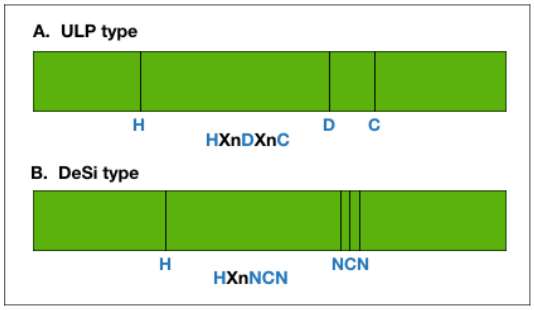
Name	Length (AA)	Subcellular Localisation	Predominant Function	SUMO Isoform Processed	SUMO Isoform Deconjugated	Biological Implication
ESD4	489	Nucleus	Isopeptidase	SUM1; SUM2	SUM1; SUM2	Control flowering time and plant development (Murtas *et al.*, 2003)
ULP1a / ELS1	502	Cytosol	Endopeptidase	SUM1; SUM2; SUM3 (weakly)	SUM1; SUM2	Control flowering time and plant development, however less impact than in *esd4* plants (Hermkes *et al.*, 2011)
ULP1b	341	Nucleus	Endopeptidase	Not Tested	Not Tested	
ULP1c / OTS2	571	Nucleus and nuclear foci	Both	SUM1; SUM2	SUM1; SUM2	Salt stress responses, DELLA-dependent regulation of growth, modulating SA signalling and de-SUMOylation of phytochrome-B (Conti *et al*., 2009; Castro *et al*., 2016)
ULP1d / OTS1	584	Nucleus	Both	SUM1; SUM2	SUM1; SUM2	Salt stress responses, DELLA-dependent regulation of growth, modulating SA signalling and de-SUMOylation of phytochrome-B (Conti *et al.*, 2009; Castro *et al.*, 2016). 35S:OTS1 had increased salt tolerance and reduced level of SUMO-conjugated proteins (Benlloch and Lois, 2018)
ULP2a / SPF2	774	Nucleus	Endopeptidase	SUM1	Yes	*spf1/spf2* double mutants exhibit severe defects in gametogenesis and embryo development (Liu *et al.*, 2017)
ULP2b / SPF1 / ASP1	963	Nucleus	Endopeptidase	SUM1	Yes	Regulates flowering time (Kong *et al.*, 2017). *spf1/spf2* double mutants exhibit severe defects in gametogenesis and embryo development (Liu *et al.*, 2017)

ULPs, responsible for SUMO maturation and deconjugation, constitute the most abundant family among members of the SUMOylation machinery and exhibit specificity for SUMO isoforms and target substrates. These sumo proteases generally consist of a conserved C-terminal domain and variable N-terminal domain. The former domain contains the conserved ULP1-Catalytic (ULP1-C) domain enclosed by a catalytic triad of histidine, or aspartic acid, and cysteine residues (Conti *et al.*, 2008). Whereas the highly divergent N-terminal domain has a significant role in regulating ULP activity *in vivo* and is thought to confer substrate specificity (Mukhopadhyay and Dasso, 2007; Hickey *et al*., 2012). This structural organisation is exhibited in yeast ULP1 and human SENTRIN-SPECIFIC PROTEASE (SENP) 1, SENP2, SENP4, and SENP5. In the instance of ULP2s however, the ULP1-C domain is found in the middle of the protein. As presented in Table 1.1, Arabidopsis ESD4 and ULP proteins possess SUMO maturation and deconjugation activities.

Recently, studies have identified a new separate class of SUMO proteases initially described in animal systems, belonging to the evolutionarily distinct C97 family of cysteine proteases; DeSis (Nayak and Müller, 2014). Unlike ESD4 and ULPs, DeSi proteins function only in the removal of SUMO from target substrates and exhibit no SUMO maturation processing activities. In addition, DeSi proteins exhibit extremely high specificity

to target substrates and are characterised by permuted papain fold peptidases of the double-stranded RNA viruses and eukaryotes (PPPDE) domains (Shin *et al.*, 2012; Nayak and Müller, 2014). DeSi-1 and DeSi-2 are DeSi family member proteins found in plants and metazoa, however, are absent in lower eukaryotes. These DeSi proteins have demonstrated their ability to deconjugate SUMO1, 2, and 3 from substrate proteins (Shin *et al*., 2012). Crystal structures of DeSi-1 confirmed the protein forms a dimer where the active site is situated between the groove formed by the two subunits. This active site region comprises of two conserved cysteine and histidine residues that make up the catalytic dyad, as shown in Figure 1.3 (Nayak and Müller, 2014).

Figure 1.3 – Schematic diagram of the ULP and DeSi SUMO protease and their characteristic catalytic motif. A) Displays the ULP SUMO protease catalytic triad represented by the distinctive Histidine-Aspartic Acid-Cysteine triad (H-D-C), with any amino acid (X) of an undetermined length (n) between each residue. B) Shows the DeSi SUMO protease catalytic motif defined by Histidine-Asparagine-Cysteine-Asparagine (H-NCN). In this case, the NCN triad is sequential with an undetermined amino acid sequence between the triad and the H.



As of present, there is limited knowledge on DeSi-1 and DeSi-2 proteins aside from data confirming their restricted substrate specificity with the transcriptional repressor BTB-ZINC FINGER EFFECTOR LYMPHOCYTES (BZEL), found in mice (Shin *et al.*, 2012). In addition to the literature, a new class of proteases related to DeSis has been found in humans; the Ubiquitin-Specific Protease-Like (USPL) protease. USPL1 is critical in the cell proliferation stage of the cell cycle and seems to have a higher affinity for SUMO2 and SUMO3 then for SUMO1 (Kolli *et al.*, 2010; Schulz *et al.*, 2012). The role of SUMO proteases in plants is discussed in the forthcoming section 3, with an overview of the current literature and investigation of DeSi proteases provided.

1.5. Aim of Research

As SUMO has been implicated in various stress-induced adaptive responses in plants, it is crucial to further elucidate the role of SUMO in stress signalling pathways, particularly in the critical machinery component that confers substrate specificity; SUMO proteases. This thesis aims to define and characterise the role of two newly identified DeSi proteases in Arabidopsis plants; *At4g25660* (AT60) and *At4g25680* (AT80) of the DeSi2 subgroup. The characteristics and properties of the two DeSi SUMO proteases were first defined using bioinformatic, genetic and proteomic analyses using both online resources and laboratory experiments. Genetic KOs and overexpressing transgenics of AT60 and AT80 proteases were generated, genotyped and phenotypically analysed in normal and stress-induced conditions to investigate the function of the DeSi proteases in Arabidopsis development and defence responses. The overexpressing transgenic lines were also examined by confocal microscopy to ascertain the organ-specific and subcellular localisation of the two DeSi proteases in normal conditions and in response to stress. Lastly, the biochemical properties of the AT60 protease were investigated using protein analyses to demonstrate and validate the protein's SUMO protease activity in cleaving conjugated SUMO from a target substrate.

2. Materials and Methods

2.1. Materials

The chemicals used in this study were bought from Sigma-Aldrich, Fisher Scientific, VWR or Melford.

2.1.1. Plant Material

Arabidopsis WT Columbia (Col-0) seeds were obtained from lab stocks. All mutants and reporter lines were in the Col-0 background and also obtained from lab stocks unless otherwise stated as indicated in Table 2.1. *Nicotiana benthamiana* (*N. benthamiana*) seeds were obtained from lab stocks and grown in-house for experiments.

Table 2.1 – Arabidopsis mutant and reporter lines

Genotype	Segregation Status	Experimental Use	Origin
desi3a-1 *AT1G47740* KO	Homozygous	Control	Inhouse (Dr. Yates)
ots1 ots2 mutants	Homozygous	Control	Inhouse (Dr. Srivastava)
AT4G25680 KO (SALK_064598C)	Homozygous	Background genotype for the double KO mutant	The European Arabidopsis Stock Centre (NASC)
461 (Col-0 + YFP)	Homozygous	YFP Control for Microscopy	Inhouse (Dr. Knight)
642 (Col-0 + Strep-His-YFP)	Homozygous	YFP Control for Microscopy	Inhouse (Dr. Knight)

2.1.2. Antibiotics

Antibiotic stocks listed in Table 2.2, were filter sterilised and aliquoted into individual Eppendorfs under a fume hood.

Table 2.2 – Antibiotics prepared and stored at -20°C

Antibiotic	Solvent	Working Concentration (µg ml^{-1})
Carbenicillin	dH^2O	100
Chloramphenicol	Ethanol	34
Gentamycin	Ethanol	10
Hygromycin	dH^2O	50
Kanamycin	dH^2O	50
Rifampicin	Methanol	12.5
Spectinomycin	dH^2O	50

2.1.3. Antibodies

Primary and secondary antibodies were prepared to their respective concentrations as indicated in Table 2.3, before use and subsequently stored at -20°C. Primary antibodies were re-used up to 5 times whereas secondary antibodies were discarded.

Table 2.3 – Primary and secondary antibodies prepared and stored at -20°C

Antibody		Host	Working Concentration (Antibody: TBST)	Supplier
Primary	Anti-HA	Rat	1:3000	Roche
	Anti-GST	Rat	1:5000	Abcam

	Anti-His	Mouse	1:5000	Roche
	Anti-SUMO1	Rabbit	1:10,000	Produced inhouse
	Anti-GFP	Rabbit	1:8000	Abcam
Secondary	Anti-Mouse-HRP	Sigma	1:20,000	Sigma
	Anti-Rat-HRP	Sigma	1:20,000	Sigma
	Anti-Rabbit-HRP	Sigma	1:20,000	Sigma

2.1.4. Vectors

Table 2.4 – Vectors used in experiments stored at -20°C

Vector		**Resistance**	**Protein Tags**	**Expression**
Entry	pENTR/D-TOPO	Kanamycin		*Escherichia coli (E. coli)*
Entry	pCBC-DT1T2	Chloramphenicol		*E. coli*
Destination	pHEE401E	Kanamycin (*E. coli*) Hygromycin (plants)	3x FLAG + NLS (N-terminus) NLS (C-terminus)	*E. coli* and Plant
Destination	pEarleyGate101 (pEG101)	Kanamycin (*E. coli*) BASTA (plants)	YFP + HA (C-terminus)	*E. coli* and Plant
Destination	pDEST 15	Ampicillin (Carbenicillin)	GST (N-terminus)	*E. coli*
Destination	pDEST 17	Ampicillin (Carbenicillin)	HIS (N-terminus)	*E. coli*

2.1.5. Media, Buffers and Solutions

Table 2.5 - Media, buffers and solutions prepared at respective compositions and used in experiments

Buffer / Solution / Media		**Composition (brought to desired volume with dH^2O)**	**Buffered pH**
Media	Liquid LB	10g L^{-1} Tryptone, 5g L^{-1} Yeast Extract, 5g L^{-1} NaCl	7.2
	Solid LB	10g L^{-1} Tryptone, 5g L^{-1} Yeast Extract, 5g L^{-1} NaCl, 1.2% w/v Agar	7.2
	MS Solid	2.2g L^{-1} MS Basal Salt Mixture, 7.8g L^{-1} Phytoagar	7.0
	MS Liquid	2.2g L^{-1} MS Basal Salt Mixture	7.0
	Super Optimal Broth with Catabolite Repression (SOC)	20g L^{-1} Tryptone, 5g L^{-1} Yeast Extract, 10mM NaCl, 2.5mM KCl, 10mM $MgCl_2$, 10mM $MgSO_4$, 20mM Glucose	7.0
Buffer	Protein Extraction Buffer	150mM NaCl, 50mM Tris-HCl (pH 8), 1% Igepal, 0.5%, 0.2% Sodium Deoxycholate, 0.1% SDS, 1mM EDTA, 50mM N-Ethylmaleimide (*NEM*), Proteinase Inhibitor Tablet	n/a
	1x Running Buffer	25mM Tris, 192mM, Glycine, 0.1% SDS	n/a
	1x Transfer Buffer	25mM Tris, 192mM Glycine, 10% Methanol	n/a
	1x Tris-buffered saline and Tween-20 (TBST)	50mM Tris (pH 7.4), 150mM NaCl, 0.1% Tween-20	n/a
	Phosphate-buffered saline (PBS)	10mM Na_2PO_4, 137mM NaCl, 2.7mM KCl	7.4
	4x Sodium Dodecycl Sulfate (SDS) Loading Buffer	200mM Tris-HCl (pH 6.8), 400mM dithiothreitol (DTT), 8% SDS, 40% Glycerol, 1% Betamercaptoethanol, 0.4% Bromophenol Blue	n/a
	His Binding Buffer	20mM Na_2PO_4, 0.5M NaCl, 40mM Imidazole	7.4
	His Elution Buffer	20mM Na_2PO_4, 0.5M NaCl, 500mM Imidazole	7.4
	Wash Buffer	20mM Na_2PO_4, 0.5M NaCl, 20mM Imidazole	7.4
	SUMO Protease Buffer	50mM Tris-HCl (pH 8.0), 150mM NaCl, 0.2% Igepal (Nonidet P-40), 1mM DTT	n/a
Solution	Blocking Solution	5% w/v Non-fat Milk Powder in 1xTBST	n/a

	Coomassie Blue Stain Reagent	0.25% w/v Brilliant Blue, 50% v/v Methanol, 10% v/v Glacial Acetic Acid	n/a
	De-staining Solution	20% Methanol, 10% Glacial Acetic Acid	n/a
	Ponceau S Stain	0.5% w/v Ponceau S, 1% v/v Glacial Acetic Acid	n/a
	ECL 'Solution A'	2.5mM Luminol, 0.4mM p-coumaric acid, 100mM Tris pH 8.5	n/a
	ECL 'Solution B'	0.02% Hydrogen Peroxide, 100mM Tris pH 8.5	n/a

To prepare the Luria-Bertani (LB) and Murashige and Skoog (MS) agar plates, the media was first autoclaved and stored in a water bath at 55°C for 20 minutes. Under a fume hood, approximately 20-25ml (for cylindrical bacterial or plant tissue culture plates) or 55-60ml (for square plant tissue culture plates) of media was poured into the respective sterilised plates. The plates were left to solidify for 30 minutes, before being seeing sealed and stored at 4°C.

2.1.6. Bacteria

All bacterial strains were grown overnight in LB liquid culture before storage. 2ml of the culture was moved into an Eppendorf and mixed with glycerol to a final concentration of 15%. The Eppendorfs were frozen in liquid nitrogen and subsequently stored at -80°C.

Table 2.6 – Bacterial strains provided and used, stored as glycerol stocks at -80°C

Organism	Strain	Resistance
E. coli	DH5α	No Selection
E. coli	BL21-CodonPlus (DE3)-RIL	Chloramphenicol
E. coli	BL21 (DE3) CodonPlus – Constituting Arabidopsis SUMOylation machinery proteins (Okada *et al.*, 2009)	Chloramphenicol and Streptomycin (Spectinomycin)
Agrobacterium tumefaciens (Agrobacterium)	GV3101, pMP90	Rifampicin and Gentamicin

The bacterial strain, *E. coli* BL21 (DE3) CodonPlus constituting Arabidopsis SUMOylation machinery proteins (Okada *et al.,* 2009) will hereafter be referred to as *E. coli* SS+ cells.

Table 2.7 – Recombinant bacterial strains used and kindly donated by in-house lab personnel, stored as glycerol stocks at -80°C

Organism and Strain	Vector	Recombinant Protein	Protein Tags	Selection	Clone Origin
E. coli SS+ cells	pDEST17	JAZ6	HIS (N-terminus)	Chloramphenicol, Streptomycin (Spectinomycin) and Carbenicillin	Inhouse (Dr. Srivastava)
E. coli DH5α	pENTR/ D-TOPO	AT60		Kanamycin	In-house (Dr. Orosa)
E. coli DH5α	pENTR/ D-TOPO	AT80		Kanamycin	In-house (Dr. Orosa)
Agrobacterium GV3101	pEG101	AT60	YFP + HA (C-terminus)	Kanamycin (*E. coli*) BASTA (Plants)	Inhouse (Dr. Orosa)
Agrobacterium GV3101	pEG101	AT80	YFP + HA (C-terminus)	Kanamycin (*E.coli*) BASTA (plants)	Inhouse (Dr. Orosa)

2.2. Methods

2.2.1. Nucleic Acid Isolation

2.2.1.1. RNA Extraction

Leaf tissue frozen in liquid nitrogen was ground extensively into a fine powder using a cooled mortar and pestle. 750µl of trizol (Zymo Research, Irvine, USA) was added into the leaf tissue powder and vortexed. The Direct-zol™ RNA miniprep kit (Zymo Research) was used to extract RNA including the In-Column DNase 1 digestion with no changes from the protocol within the kit. RNA concentration was measured using a NanoDrop™ One microvolume UV-Vis spectrophotometer (Thermo Scientific™). The RNA was stored at -80°C.

2.2.1.2. cDNA Synthesis

2-4µg of RNA was mixed with sterile distilled water to make up a final volume of 10µl. For amplification, 1µl of oligo dT (10mM) (VWR, Radnor, USA) was added to the RNA mixture and heated at 65°C for 5 minutes then placed on ice. 4µl of 5x strand buffer (Invitrogen, Grand Island, USA), 2µl of DTT (Invitrogen), 1µl of 10mM deoxynucleotide triphosphate (dNTP) (VWR) and 1µl of RNAse out (Invitrogen) was incorporated into the RNA mixture and subsequently heated to 42°C for 2 minutes before the final addition of 1µl of Superscript II (Invitrogen). This mixture was then heated for 50 minutes at 42°C followed by 70°C for 15 minutes.

2.2.1.3. Genomic DNA Extraction

For genomic DNA Extraction, the protocol by Edwards *et al*. (1991) was followed with few modifications. Leaf tissue was first placed in an Eppendorf, frozen in liquid nitrogen and ground into a fine powder using a micro-pestle. 400µl of Edwards extraction buffer composed of 200mM Tris-HCl (pH 7.5), 250mM NaCl, 25mM EDTA and 0.5% SDS was added to the sample and vortexed for 5 seconds. This sample was then centrifuged at 13,000 revolutions per minute (rpm) for 4 minutes and 300µl of the supernatant was extracted. This supernatant was mixed with 300µl of isopropanol to precipitate the DNA and left at room temperature for 2 minutes before being centrifuged at 13,000rpm for 5 minutes. The supernatant was discarded, 200µl of 70% ethanol was added to wash away the salts and the sample was subsequently centrifuged for 5 minutes at 13,000rpm. The ethanol was removed by tapping the Eppendorf upside down leaving a pellet, which was left to dry on the lab bench for over 2 hours. The DNA pellet was then resuspended in 30µl sterile water and stored at 20°C.

2.2.2. Polymerase Chain Reaction (PCR)

2.2.2.1. Primer Design

Primers were designed using Primer-BLAST (http://www.ncbi.nlm.nih.gov/tools/primer-blast/), analysed using Integrated DNA Technologies (IDT) OligoAnalyzer (https://eu.idtdna.com/calc/analyzer) and synthesised by IDT (https://eu.idtdna.com/pages). qPCR primers were designed with Primer3

(http://bioinfo.ut.ee/primer3/) and NCBI primer-blast (https://www.ncbi.nlm.nih.gov/tools/primer-blast/). The full list of primers used throughout this study can be found in Appendix Table 8.1.

2.2.2.2. Taq Polymerase PCR

For a standard PCR, the following 10µl reaction mix (Table 2.8) was made up per reaction using 2x MyTaq™ Red Mix (Bioline).

Table 2.8 – 10µl reaction mix per reaction for a standard PCR

Component	Volume per Reaction
2x MyTaq™ Red Mix (Bioline)	5µl
Sterile Distilled Water	To 10µl
Forward Primer (10µM)	0.5µl
Reverse Primer (10µM)	0.5µl
Template DNA	1µl

This PCR reaction mix was run in a Veriti™ Thermal cycler under the conditions listed in Table 2.9.

Table 2.9 – Program for standard PCR using Veriti™ Thermal cycler

PCR Steps	Temperature (°C)	Time	Number of Cycles
Initial Denaturation	95	5 minutes	-
Denaturation	95	30 seconds	25-35 Cycles
Annealing	45-65 (depends on primer melting temperature)	30 seconds	
Extension	72	1 minute per 1 kb of gene to be amplified	
Final Extension	72	5 minutes	-
Hold	12	∞	-

2.2.2.3. Colony PCR

Colony PCRs were conducted using the same reaction mix and program conditions as the Taq Polymerase PCR with one amendment. The template DNA was obtained from an individual colony of bacteria and suspended in 30µl of sterile distilled water. 1µl of this was used as the DNA template in each reaction.

2.2.2.4. Q5 Polymerase Proof-Reading PCR

For a PCR using the Q5® High-Fidelity DNA Polymerase, the following 50µl reaction mix listed in Table 2.10 was used and performed in a Veriti™ Thermal cycler under the program conditions in Table 2.11. To see PCR results, 5µl of 10x DNA loading dye was added to PCR products, which were run on an agarose gel (gel electrophoresis discussed below).

Table 2.10 – 50µl reaction mix per reaction for Q5 PCR

Components	Volume Per Reaction
5x Q5 Reaction Buffer (NEB)	10µl
dNTPs (10mM)	1µl
Forward Primer (10µM)	2.5µl
Reverse Primer (10µM)	2.5µl
cDNA (< 1,000ng)	2.5µl
Q5 High-Fidelity DNA polymerase (NEB)	0.5µl
Sterile Distilled Water	Up to 50µl

Table 2.11 – Program for Q5 PCR using Veriti™ Thermal cycler

PCR Steps	Temperature (°C)	Time	Number of Cycles
Initial Denaturation	98	30 seconds	-
Denaturation	98	10 seconds	30 Cycles
Annealing	50–72 (depends on primer melting temperature)	30 seconds	
Extension	72	30 seconds per 1 kb of gene to be amplified	
Final Extension	72	2 minutes	-
Hold	12	∞	-

2.2.2.5. Site-Directed Mutagenesis

For a PCR reaction for site-directed mutagenesis, the following 40µl reaction mix (Table 2.12) was used and performed in a Veriti™ Thermal cycler under the program conditions in Table 2.13.

Table 2.12 – 40µl reaction mix per reaction for site-directed mutagenesis PCR reaction

Components	Volume Per Reaction
5x Q5 Reaction Buffer (NEB)	8µl
dNTPs (10mM)	0.8µl
Forward Primer (10µM)	2µl
Reverse Primer (10µM)	2µl
Plasmid (< 1,000ng)	2.5µl
Q5 High-Fidelity DNA polymerase (NEB)	0.4µl
Sterile Distilled Water	Up to 40µl

Table 2.13 – Program for site-directed mutagenesis PCR reaction using Veriti™ Thermal cycler

PCR Steps	Temperature (°C)	Time	Number of Cycles
Initial Denaturation	98	30 seconds	-
Denaturation	98	10 seconds	20 Cycles
Annealing	50–72 (Depends on primer melting temperature)	30 seconds	
Extension	72	30 seconds per 1 kb of gene to be amplified	
Final Extension	72	5 minutes	-
Hold	12	∞	-

To digest the methylated DNA template, 1µl of DpnI (NEB) was added directly to the PCR reaction, vortexed and spun down briefly before being incubated at 37°C for 2 hours. The reaction was heated at 80°C for 20 minutes to deactivate the DpnI enzyme and the mutated plasmid was transformed into DH5α *E. coli* cells.

2.2.2.6. Quantitative Real-Time PCR

For real-time PCR analyses, the universal SYBR Green Quantitative PCR protocol (Sigma) was used with few modifications. 10µl reactions consisting of components listed in Table 2.14 were run in a Rotor-Gene Q Machine (QIAGEN®) under the program conditions stated in Table 2.15.

Table 2.14 – 10µl reaction mix per reaction for real-time PCR

Components	Volume Per Reaction
Brilliant II SYBR® Green QPCR Master Mix	5µl
Forward Primer (10µM)	0.5µl
Reverse Primer (10µM)	0.5µl
Template cDNA	0.5µl

Sterile Distilled Water	7µl

Table 2.15 – Program for real-time PCR using Rotor-Gene Q Machine (QIAGEN®)

Temperature (°C)	Time	Number of Cycles
95	20 seconds	-
94	10 seconds	40 Cycles
55-65 (depends on primer melting temperature)	30 seconds	
72	2 minutes	

The actin gene (*AT2G37620*) was used as a reference gene for normalisation in all reverse transcription (RT)-PCR analyses. Technical repeats were conducted in triplicates for each sample and comparisons were performed using the $2^{-\Delta\Delta CT}$ method (Ramakers *et al*., 2003).

2.2.3. Gel Electrophoresis

Following PCR reactions, samples were separated by size using gel electrophoresis to identify PCR products. Gels were set with 0.8-1.2% agarose (Severn Biotech), with the higher concentrations allowing for better separation of smaller fragments. 1x TAE buffer (Biorad) was added to 0.8-1.2g of agarose (Severn Biotech) and heated in a microwave until the agarose fully dissolved. Per 100ml of solution, 1µl of ethidium bromide (VWR) for DNA visualisation, was added once the mixture had cooled. The gels were left to solidify, before being submerged in gel electrophoresis tanks containing 1x TAE buffer. Wells were loaded with 9µl of the PCR product and 3µl of the hyperladder, which was either 50 base pairs (bp) or 1Kb depending on the size of the fragment (Bioline) and run at ~100 volts. DNA fragments were visualised using a Ultraviolet (UV) transilluminator.

2.2.4. Gel Extraction

To extract DNA fragments from the agarose gel, the UV transilluminator was used to visualise and excise the relevant band using a scalpel blade, before being transferred to a pre-weighed 1.5ml Eppendorf. This Eppendorf was reweighed to determine the weight of the excised gel fragment. The gel extraction was then performed following instructions provided in the ZymocleanTM Gel DNA Recovery kit (Zymo Research). 30µl of water was used for the final elution step.

2.2.5. Cloning

2.2.5.1. Golden Gate Reaction

The Golden Gate cloning method was used for the CRISPR protocol. This entails initially using the cloning vector, pCBC-DT1T2, as a template to amplify a specifically designed PCR fragment. This insert was then extracted and purified, and together with the destination vector, pHEE401E, used to set up a restriction-ligation reaction using *Bsa*1 (NEB) and T4 DNA ligase (NEB) as indicated in Table 2.16.

Table 2.16 – Components for the Golden Gate reaction

Components	Volume (µl)
Purified PCR fragment (100ng/µl)	4

pHEE401E (100ng/µl)	2
10x T4 DNA Ligase Buffer (NEB)	1.5
10x BSA	1.5
BsaI (NEB)	1
T4 DNA Ligase (HC, NEB)	1
ddH^2O	4

This reaction was incubated for 5 hours in a 37°C room, 5 minutes in a 50°C water bath and 10 minutes on an 80°C heat block. The CRISPR protocol is further explained in detail in section 3.

2.2.5.2. Miniprep

The bacterial colony of interest was grown overnight at 37°C (*E. coli*) or 28°C (*Agrobacterium*) in a 10ml universal flask containing the appropriate antibiotics (listed in Table 2.6). The culture was spun down by centrifugation at 5000rpm for 10 minutes at 4°C. The supernatant was disposed and from the pellet, the plasmid was purified using the QIAprepR spin miniprep kit (Qiagen) following the provided protocol. The concentration of the plasmid was measured using a NanoDrop™ (Thermo Scientific™).

2.2.5.3. LR Reaction (into gateway donor vector)

The LR Reaction was used to transfer genes of interest from the entry vector into the appropriate destination vector. 50-150ng of the pENTR/D-TOPO vector and 150ng of the destination vector were mixed with 1µl of TE buffer (pH 8.0) and 0.5µl of LR Clonase II enzyme (Invitrogen). This solution was spun down and incubated at room temperature for 1.5 hours. The proteinase K solution (Invitrogen) was added to inhibit the reaction and incubated at 37°C for 10 minutes. This solution was then transformed into *E. coli* cells.

2.2.6. Transformation

2.2.6.1. *E. coli* (DH5α and BL21)

200µl aliquots of chemically competent *E. coli* cells (DH5α and BL21) in Eppendorfs were thawed on ice. 1-5µl of the DNA product (plasmid from D-TOPO, LR reactions or golden gate reactions) was added to the competent cells and the Eppendorf was flicked 3-5 times to mix the components. The Eppendorf was then transferred onto ice for 5 minutes, dropped into liquid nitrogen for 5 minutes and then placed in a 42°C water bath before being placed on ice again. 1ml of SOC media was added to each Eppendorf and left on a shaker at 200rpm for 1 hour in 37°C. Eppendorfs were then centrifuged for 1 minute at 13,300rpm to separate the cell pellet. 100-200µl of the supernatant was saved whilst the rest was discarded. The preserved supernatant was used to resuspend the cell pellet. Subsequently, the cell solution was spread onto LB agar plates containing the appropriate antibiotics for the vector and left to incubate for 16 hours in 37°C conditions. Colonies on the LB agar plates were selected for a colony PCR to confirm the transformation was successful.

2.2.6.2. Agrobacterium (GV3101)

Following a similar protocol to the *E. coli* transformation method, 200µl aliquots of chemically competent *Agrobacterium* cells were thawed on ice prior to the addition of 1-5µl of DNA product. This was incubated for 5 minutes on ice, dropped in liquid nitrogen for 5 minutes and then transferred into a 37°C water bath before being placed on ice again. 1ml of LB liquid media was added before being left on a shaker for 2 hours in 28°C.

The cells were spun down and re-suspended in 100-200µl of the retained supernatant before the cell solution was spread onto LB agar plates containing the appropriate antibiotics for the vector and left to incubate for 48 hours in 28°C. Colonies on the LB agar plates were selected for a colony PCR.

2.2.7. Protein Analysis

2.2.7.1. Protein Expression

To express the recombinant proteins in *E. coli* for DH5α, BL21 and SS+ strains, the transgenic *E. coli* cells were grown in 10ml liquid LB cultures containing the appropriate antibiotic(s) on a shaker at 200rpm overnight at 37°C. Expression profiling of the recombinant protein was then obtained by testing optimum conditions using the following protocol. 500µl of the overnight culture was added to 50ml of liquid LB within a 250ml conical flask with the appropriate antibiotics. The flask was left on a 200rpm shaker at 37°C until the optical density (OD) at 600 wavelength of the culture reached 0.6-0.8. At this point, two 1ml samples were retrieved and preserved on ice. 1mM of isopropylthiogalactoside (IPTG) (Fisher Scientific) was then added to the culture to induce protein expression from the inducible vector (pDEST15/pDEST17). Culture samples were then collected after 1, 2 and 3 hours following IPTG induction. The OD_{600} measured at each time point determined the volume of the sample collected, to ensure all samples contained the same number of bacterial cells as the 1ml pre-induced sample. All collected samples were then centrifuged at 13,300rpm for 1 minute at 4°C and the supernatant was discarded leaving just the cell pellet. To split the samples into soluble and insoluble fractions, a stock of 400µl of BugBuster (Novagen, Billerica, USA) and 40µl of 100mM Phenylmethylsulfonyl fluoride (PMSF) was made. 100µl of this mixture was aliquoted into each cell pellet sample and left to incubate at room temperature for 20 minutes. The samples were then centrifuged for 3 minutes at 13,300rpm set at 4°C. 60µl of the supernatant (soluble fraction) was retrieved and transferred into a new Eppendorf, where 20µl of 4x SDS loading buffer was added before being heated at 98°C for 3 minutes. The remaining pellet (insoluble fraction) was re-suspended in 60µl of PBS and 20µl of 4x SDS loading buffer and subsequently heated for 3 minutes at 98°C. To analyse protein content, 10-20µl samples of each fraction at each time point was then loaded onto SDS-polyacrylamide gel electrophoresis (SDS-PAGE). For visualisation and to identify optimal conditions for recombinant protein expression, Coomassie staining and immunoblotting was performed. The optimal conditions drawn from this protocol were repeated for purification. In some cases, to improve protein folding, adjustments were made to this procedure including reducing IPTG concentration and inducing cultures at a lower temperature for a longer period of time.

2.2.7.2. SDS-PAGE Gel

SDS-PAGE is an electrophoresis method that separates protein by mass and is used to analyse protein content. The Min-Protean Tetra Cell system (Bio-Rad, Hercules, USA) was used for all experiments. Hand-cast polyacrylamide gels were made with the following composition listed in Table 2.17.

Table 2.17 – Composition of the resolving and stacking gel. Acrylamide percentage varies depending on the molecular weight of the proteins.

Gel Type	Composition
Resolving Gel	10-15% acrylamide
	0.375M Tris pH 8.8

	0.1% SDS
	0.1% ammonium persulphate (APS)
	0.04% TEMED
	Sterile H_20 up to Desired Volume
Stacking Gel	5% acrylamide
	0.125M Tris pH 6.8
	0.1% SDS
	0.1% APS
	0.01% TEMED
	Sterile H_20 up to Desired Volume

When making the gels, the resolving gel was initially poured into a 15mm gel mould and 100% isopropanol (Fischer Scientific) was added on top to create a levelled gel. The gel was left to set for 30 minutes before the isopropanol was discarded. The stacking gel was pipetted onto the resolving gel and a 10 or 15 well comb was placed into the mould. Once the gel had set, the comb was removed, and the gel was placed into the conductor unit of the gel tank where it was subsequently submerged in 1x running buffer. To denature and dilute the sample, 4xSDS PAGE loading buffer was added to the samples at a 1:3 ratio and heated for 5 minutes at 98°C. With the first lane loaded with 5µl of PAGE ruler protein ladder (Thermo Scientific), protein samples were loaded into the remaining wells. Electrophoresis was run at 80-100V until the loading dye reached the bottom of the gel. To visualise the protein bands, Coomassie staining or western blotting was performed.

2.2.7.3. Coomassie Staining

Following SDS-PAGE, the resolving gel was left to stain in Coomassie dye on a horizontal shaker for 45 minutes. The gel was then left to soak in de-staining solution on a horizontal shaker until the protein bands become clearly visible.

2.2.7.4. Western Blotting

Western blots were performed to visualise protein bands in transient expression assays in *N. benthamiana* leaves, transgenic Arabidopsis lines and recombinant E. coli strains. The polyvinylidene fluoride (PVDF) membrane was first cut to the size of the resolving gel, activated in 100% methanol and soaked in pre-chilled 1x transfer buffer. All components involved in the transfer process were submerged in 1x transfer buffer. This included, from the back side of the clamp ready cassette: a sponge, 2 sheets of membrane-sized blotting paper, the resolving gel, the activated membrane, 2 sheets of blotting paper and a sponge. Once all components were sandwiched tightly between the cassette, it was fitted into a conductor unit with an ice pack and the gel tank was filled with 1x transfer buffer. This was left to run at 30V for 16 hours in 4°C conditions. Following the transfer, the membrane was incubated in blocking solution for over 1 hour and 30 minutes at room temperature. To remove any excess milk from the membrane, 1x TBST was used to rinse the membrane, which was subsequently incubated in the appropriate primary antibody (concentrations listed in Table 2.3) for 1-2 hours depending on the antibody. A series of 5x 5minute membrane washes in 1x TBST was then performed on a fast shaking rocker. The relevant second antibody (concentrations listed in Table 2.3) was then added to the membrane for 1 hour of incubation. A series of 5x 5minute 1x TBST washes proceeded the incubation on a high-speed shaker to wash away any unbound or non-specifically bound antibodies. The ECL solution composed of ECL solution A and B at a ratio of 1:1, was prepared and the solution (2ml per membrane) was poured over the membrane for 1 minute. To remove excess ECL solution, the membrane was lightly dried on

a paper towel before being laid in between two transparency films within a light-proof cassette. The cassette was only opened in the dark room where Fujifilm X-ray film (Fisher Scientific) was exposed to the membrane for different lengths of time depending on the antibody. The film was subsequently developed using a Xograph Compact 4x Automated Processor (Xograph Imaging Systems) to visualise the protein bands.

2.2.7.5. Protein Extraction from *N. benthamiana* or Arabidopsis Leaves

1g of leaf tissue was collected from the leaves of *N. benthamiana* or Arabidopsis plants and frozen in liquid nitrogen. The tissue sample was ground to powder form using a mortar and pestle prior to the addition of 2µl g^{-1} of protein extraction buffer. The mixture was again ground until it reached a thick consistency. For *N. benthamiana* samples, 1.5 w/v of PVPP (Sigma Aldrich) was also added and mixed with the sample to inhibit any phenolics in the plant sample. The mixture was subsequently centrifuged at 10,000rpm for 8 minutes at 4°C. The resulting supernatant was retrieved and centrifuged again under the same conditions. The re-isolated supernatant was then diluted with 4x SDS loading buffer before being heated at 98°C for 5 minutes, so it could either be stored at -20°C or loaded onto an SDS-PAGE gel for protein separation. To measure the protein concentration of the sample, a Direct Detect Spectrometer (Merck Milipore) was used.

2.2.7.6. Protein Extraction for Purification

For protein purification, bacterial cultures were scaled up to 1L using the starter overnight culture and selective antibiotics. The cultures were subsequently induced using optimal conditions defined during expression profiling once the OD600 of the culture reached 0.6-0.8. The BugBuster solution was then prepared comprising of 8ml BugBuster, 800µl PMSF and if the recombinant protein to be expressed was not a protease protein, 1 tablet of cOmplete™ Mini EDTA-Free Protease Inhibitor (Roche, Indianapolis, USA). The induced bacterial cultures were then centrifuged at 5000rpm for 20 minutes at 4°C and the supernatant was discarded. The pellets were weighed and 2.5ml of the BugBuster solution per 1 gram of the pellet was added and incubated at room temperature on a shaker for 30 minutes or until the pellet was fully solubilised. This mixture was then centrifuged for 10 minutes at 5000rpm in 4°C. The supernatant was centrifuged once again in the same conditions to filter out any solid debris. The supernatant was subsequently passed through a 0.45µl filter before being stored on ice until undergoing following purification experiments.

2.2.7.7. Fast Protein Liquid Chromatography (FPLC)

To purify proteins with a histidine (His) tag, the HisTrap (GE healthcare, Little Chalfont, USA) column and AKTA machine was used. The 1ml HisTrap column was first washed with 5ml of deionised water and then equilibrated by 10ml of binding buffer at a 2ml min^{-1} flow rate. The filtered soluble supernatant was then added to the column at a slower rate of 1ml min^{-1} and the flow through was collected. 15ml of binding buffer at a 1ml min^{-1} flow rate was added to the column to wash it and the final 1ml of flow through was collected. Finally, 10ml of the elution buffer at the same flow rate was passed through the column and 1ml elution fractions were collected. UV absorbance levels of the sample flowing through the column were measured by the AKTA machine providing information on protein concentration of each elution fraction. The columns were subsequently washed and stored. 20µl of 4xSDS loading buffer was then added to the flow through, wash and eluted samples and run on SDS-PAGE to check for protein content. Original fractions were stored at -80°C.

2.2.7.8. *In vitro* deSUMOylation assay

To analyse the SUMO protease activity of AT60 in cleaving SUMO from the target substrate JAZ6, both AT60 and JAZ6 protein elutes had to be concentrated and dialysed from their purification fractions. This was performed using the Amicon® Ultra-0.5 centrifugal filter devices at 4°C. 1K cut off columns were used to concentrate 1ml purified protein fractions to a 250µl volume. The concentrated purified proteins then underwent 4 repetitions of dialysis washes with the SUMO protease buffer and again concentrated down to 250µl. Direct Detect® infrared spectrometer (Merck) was used to measure protein concentration. A reaction volume at a minimum of 100µl containing a ratio of 5:1 JAZ6 to AT60 was then established. This reaction was left overnight at 28°C. 4xSDS loading buffer was then added to the reaction and heated before being run on SDS-PAGE.

2.2.8. Plant Growth and Treatment

2.2.8.1. Arabidopsis Seed Sterilisation for Tissue Culture

1.5ml Eppendorfs with 15mg of seeds were placed in an air-tight box within a fume hood cabinet. 97ml of hypochlorite was first added to a conical flask and placed inside the box. 3ml of hydrochloric acid was then added to the hypochlorite and the lid of the box was immediately shut tight and left overnight. The beaker was then removed from the box before the seeds within the Eppendorfs were moved to a sterile laminar flow cabinet for at least 2 hours for airing.

2.2.8.2. Arabidopsis Growth Conditions

Sterilised seeds were distributed on MS plates under a sterile laminar flow cabinet and then transferred to a 4°C room for 72 hours for stratification. Seedlings were then moved to a Plant Growth Chamber set to long day conditions, which was programmed for 16 hours of light at 22°C and 8 hours of dark at 20°C. Plants were then grown for 21 days on MS plates before being transferred to wet Levington F2 plus sand compost soil.

2.2.8.3. Floral Dipping for *Agrobacterium*-mediated Transformation of Arabidopsis

Successfully transformed *Agrobacterium* containing recombinant vectors were used to transform Arabidopsis plants. 10ml overnight cultures of LB, recombinant *Agrobacterium* and the appropriate antibiotics were used to inoculate 1L of LB containing the same antibiotics and grown overnight on a shaker at 28°C. The following day, the bacterial cultures were spun down at 4500rpm for 10 minutes and the supernatant was discarded. The pellet was then re-suspended in 1L of 5% sucrose. 200µl of Silwet L-77 was then added to the medium. The flowering Arabidopsis plants were prepared for floral dipping by trimming off all flowers and siliques. The plants were placed upside down in the dipping medium for 30 seconds and then placed sideways on a tray and covered with a plastic bag overnight. The next day, the plastic bags were removed, and the plants were stood upright and grown in normal conditions.

2.2.8.4. Transient Expression in *N. benthamiana*

For transient assays in *N. benthamiana* plants, the gene to be expressed was cloned into a pEG series vector containing a yellow fluorescent protein (YFP) C-terminal fusion (pEG101) and transformed into *Agrobacterium* cells. 10ml LB cultures containing this recombinant *Agrobacterium* strain and P19 (RNAi silencing inhibitor) with appropriate antibiotics were grown overnight on a shaker at 28°C. The culture was then centrifuged at 5000rpm for 5 minutes at 4°C and the supernatant was discarded. The pellet was then re-suspended in 10mM $MgCl_2$ to reach an OD_{600} of 0.4. Following this, 0.1mM of Acetosyrine was added to the solution. A mixture composed of P19 and the recombinant bacterial solution at a 50:50 ratio was made. This mixture was then infiltrated into the leaves of *N. benthamiana* with a 1ml syringe and the plants were watered and left for 2-3 days in normal growth conditions. To visualise the protein, a small 0.5cm^2 section of the infiltrated *N. benthamiana* leaves were extracted and visualised on a Leica TCS SP5 confocal laser scanning microscope (LSM).

2.2.9. Arabidopsis Phenotyping Assays

2.2.9.1. Root Length Assay

Seeds were first sterilised and sown onto MS plates before being transferred to dark conditions at 4°C for 72 hours. The seeds were then grown in long day conditions for 4 days and then transferred to either control (MS media) or treatment plates for 5 days. Treatment plates consisted of ABA (25µM) as an abiotic elicitor and the flagellin peptide fragment, flg22 (250nM) as a pathogen reponse elicitor. The seedlings were grown on the plates for 5 days in long day conditions and photographed each day next to a ruler to normalise the scale. The pictures of the seedlings and their root lengths were analysed using the Fiji software. Root elongation measurements were then analysed statistically using the non-parametric Mann-Whitney U test when comparing between 2 samples and the non-parametric Kruskal-Wallis test when comparing between 3 or more samples.

2.2.9.2. Fresh Weight Assay

The protocol described in the root length assay was followed. After the final day photographs of the seedlings were taken, all tested seedlings were initially dried on a paper towel before being weighed on an electronic balance. The Mann-Whitney U test and Kruskal Wallis test were used to statistically analyse the fresh weight measurements for each genotype grown in normal and stress-induced conditions.

2.2.10. Microscopy

2.2.10.1. Confocal Microscopy

For the visualisation of transient expression assays in *N. bethamiana* leaves, approximate 0.5cm^2 sections of the infiltrated *N. bethamiana* leaves were extracted and mounted on a microscope slide (Fischer Scientific) with dH_2O and a 22x22mm coverslip (Menzel-Glaser, Waltham, USA). The slide was placed on the stage of a Leica SP5 LSM microscope (Leica, Berlin, Germany) and viewed using either an x40 or x64 objective oil lens. To excite YFP-tagged proteins, the Argon laser at 514nm was used.

2.2.10.2. Compound Light Microscopy

For subcellular protein localisation of overexpressing Arabidopsis seedlings, a Zeiss LSM 880 confocal microscope with an x40 objective oil lens was used. 4-day old transgenic seedlings were mounted onto a microscope slide with dH_2O and a 22x22mm coverslip and placed on the stage of the Zeiss LSM 880 microscope. As the recombinant proteins were tagged with YFP, the Argon laser at 514nm was used.

2.2.10.3. Analysis of Confocal Images

Images were initially processed with the Leica Application Suite Advanced Fluorescence (LAS AF) Lite software (v2.63 build 8173). For assessing changes in fluorescence, Fiji was used to measure the total mean fluorescence of YFP-tagged proteins. A polygon selection was used to draw around all sections of each tissue per image measuring the area, integrated density and mean grey value. The average of each value was calculated for each tissue. Background fluorescence was measured in areas excluding the cellular structures with no fluorescence and the mean fluorescence of background readings was calculated. The mean area of the selected tissue was multiplied by the mean fluorescence of background readings, which was then subtracted from the mean integrated density. This was repeated for each tissue per image and measurements were taken on over 5 images of each section from at least 3 different individual seedlings per genotype. The average was then taken from all analysed images per tissue.

2.2.10.4. Stress Response Assays

Stress response assays were conducted to visualise the localisation and expression of the overexpressed protein in response to stress in the overexpressing Arabidopsis seedlings. The phytohormone, ABA and pathogen response elicitor, flg22 were used as stress inducers for this experiment. 4-day old seedlings were transferred to liquid MS media in a Petri dish and 10μM of ABA was added to the Petri dish and left on a shaker for 2 hours before being immediately mounted on to a microscope slide to be visualised. For flg22, 200nM was added to the Petri dish and left on a shaker for 30 minutes before being visualised under the Zeiss 880 microscope.

3. Investigating the DeSi SUMO Proteases and the Generation and Analysis of AT60-AT80 Double KO Mutants

3.1. Introduction

As mentioned in section 1, SUMO proteases are critical in the SUMO pathway and functionally diverge from one another attributing to their potential role in substrate specificity (Table 1.1). Following the recent discovery of the DeSi-type proteases, there has been significant work conducted on further elucidating this class of SUMO proteases and in particular, uncovering their presence and function in plant systems. Research in this area would provide a greater understanding of the role and significance of SUMO proteases in the SUMO system, especially if they are implicated in the developmental or stress signalling pathways in plants.

In very recent years, bioinformatic analyses used the conserved catalytic motif of DeSi-type proteases to search the Arabidopsis proteome (Orosa *et al*., 2018). This research subsequently led to the discovery of eight new putative DeSi proteins. Primary structure analyses of these newly found proteins revealed all eight DeSi-like proteins have conserved regions including the DeSi catalytic motif aligning to each other and to the human DeSi-1 protein (Figure 3.1).

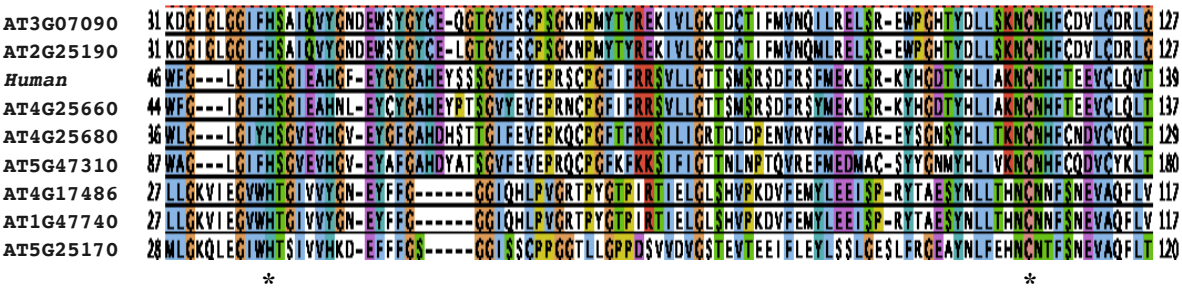


Figure 3.1 – Amino acid sequence alignment of the catalytic domain of the eight Arabidopsis DeSi-type SUMO proteases with the human DeSi-1 protease. The sequences show areas of homology and variation. Colouring intensity reflects the degree of amino acid conservation. Asterisks denote the conserved DeSi catalytic motif; H-NCN. Protein sequences were retrieved from The Arabidopsis Information Resource (TAIR) database, aligned used Clustal Omega and visualised in Jalview where Clustal-based colour scheme, which colour-codes amino acids by residue type, was applied.

Furthermore, phylogenetic analyses of the eight Arabidopsis DeSi proteases revealed the proteins could be divided into three sub-groups; DeSi1, DeSi2 and DeSi3 (Figure 3.2). The sub-group DeSi3, comprises a larger set of proteins relative to the other two groups. The tree in Figure 3.2 reveals that a common ancestral protein split into two lineages; one branch leading to DeSi1 and the other to DeSi2 and DeSi3.

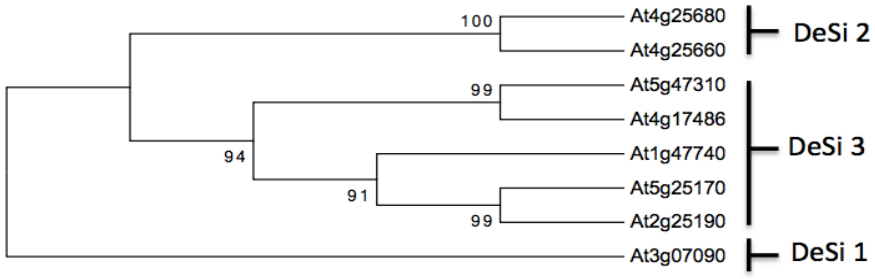


Figure 3.2 – Molecular phylogenetic analyses of Arabidopsis DeSi proteins reveals 3 distinct groups. The analysis was performed using the Maximum Likelihood method. Evolutionary history inferred based on the JTT frequency model (Jones *et al*., 1992). Numbers shown at nodes denote bootstrap values based on 1000 replicates. The analysis involved eight amino acid sequences. Figure adapted from Yates (2018).

A recent study by Orosa *et al*. (2018), further characterised one of the eight DeSi proteins; *AT1G47740* (referred to as Desi3A from here onwards). Desi3A was shown to be membrane-localised and act as a key negative regulator in the immune response pathway. *desi3a-1* mutants exhibited greater resistance to pathogen infections with a heightened response in detecting pathogenic elicitor molecules. The findings from this study concluded that flagellin perception induces the degradation of Desi3A, which promotes SUMOylation and subsequently activation of the key receptor for the perception of bacterial pathogens, FLAGELLIN SENSING 2 (FLS2), consequently triggering intracellular immune signalling.

As this study proved the vital role Desi3A plays in pathogen resistance and innate plant immunity, it further highlights the need to elucidate and characterise the newly identified DeSi SUMO proteases. In order to provide in-depth analyses on these putative DeSi proteins, this study focused only on the DeSi2 sub-group. The aim of this chapter was to define the characteristics and functional of the two Arabidopsis DeSi proteases comprising the DeSi2 sub-group; *At4g25660* (AT60) and *At4g25680* (AT80).

3.2. Investigating the DeSi SUMO proteases

3.2.1. Bioinformatics and proteomic analyses of the two DeSi SUMO proteases

The AT60 and AT80 DeSi SUMO proteases were selected for this study. Before working on the selected proteins, online resources and software programs were used to initially find out more information regarding the proteases biochemically, genetically and phenotypically. The AT60 protease is of 255 amino acids in length and has the molecular weight of 28.16 kDa. The AT80 protein is of 252 amino acids in length and weighs 27.8 kDa. A sequence alignment analysis of the two proteins revealed they share 93% identical amino acids with only 1.2% of the aligned protein sequences consisting of gaps, where they share no identical or similar amino acids. The sequence alignment analysis scored an E-value of 0, indicating the results were highly accurate. Therefore, the analysis reported in Figure 3.3, suggests strong evidence that the AT60 and AT80 proteases are a result of gene duplication events with a strong chance of genetic redundancy.

Figure 3.3 – Amino acid sequence alignment of AT60 and AT80 DeSi SUMO proteases. The sequences display greater areas of homology and minimal variation. Colouring intensity reflects the degree of amino acid conservation. Protein sequences were retrieved from TAIR database, aligned used Clustal Omega and visualised in Jalview where the Clustal-based colour scheme, which colour-codes amino acids by residue type, was applied.

To further validate the similarity in protein structure between AT60 and AT80, computational 3D modelling using the online visual analytic tool, ePlant browser, was carried out on both DeSi proteins. ePlant browser is a software that connects to various publicly available web services to download genome, proteome, interactome, transcriptome, and 3D molecular structure data for one or more genes of interest (Waese *et al.*, 2017). For 3D modelling, the ePlant browser uses the Phyre2 web portal for protein prediction, modelling and analysis. The predicted 3D structures of both proteins are displayed in Figure 3.4. The predicted 3D structures of AT60 and AT80 proteins are very similar to one another. Both proteins have a particular shape and structural formation to the β-sheets that cross over one another with the sheets spiralling in opposite directions. The structure of the β-sheets between AT60 and AT80 are close to identical, however, AT80 appears to possess two additional β-sheets within the N-terminus region of the protein (bottom section of the protein when viewed from the angle shown in Figure 3.4). The structure and formation of the α-helices also resemble one another. The catalytic cysteine residues within the DeSi motif highlighted in light green, are situated in the exact same location in relation to the α-helices for both the AT60 and AT80 protein. Therefore, this analysis further substantiates the theory that the proteins are a result of gene duplication events.

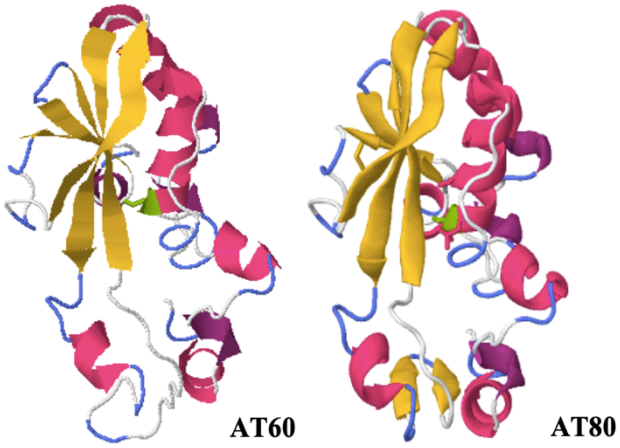


Figure 3.4 – Predicted 3D structure of AT60 and AT80 displaying striking similarities to one another. Protein sequences were obtained from TAIR database and the Phyre2 web portal on the ePlant Browser was used for protein modelling, prediction and analysis. The β-sheets are highlighted in yellow and the alpha helices in pink. The catalytic cysteine residues positioned at 115 are highlighted in light green.

To visualise and understand the spatial expression patterns and gene expression levels of the two DeSi proteases in Arabidopsis, the online ePlant browser tool was used again. This tool visually displayed the pattern and level of gene expression of AT60 and AT80 within the Arabidopsis plant under normal conditions across different plant organs and developmental stages (Figure 3.5 top panel). Using the comparative mode, the software was also able to compare the gene expression levels and patterns between the AT60 and AT80 proteases (Figure 3.5 bottom panel). This was useful for determining both the similarities of the protein's expression patterns and levels, but also highlighting their significant differences from one another. Overall, the spatial pattern and expression levels of the respective genes look similar to one another. However, in the comparative mode, it can be deduced that the AT60 expression was significantly higher in the cauline and rosette leaves relative to the gene expression level of AT80, whereas AT80 displays a more ubiquitous expression pattern across the Arabidopsis plant. Interestingly, the only plant organ in which the AT80 was more abundantly expressed was in the mature pollen of the Arabidopsis plant. Subsequent experiments were therefore performed to test the findings obtained from this analysis.

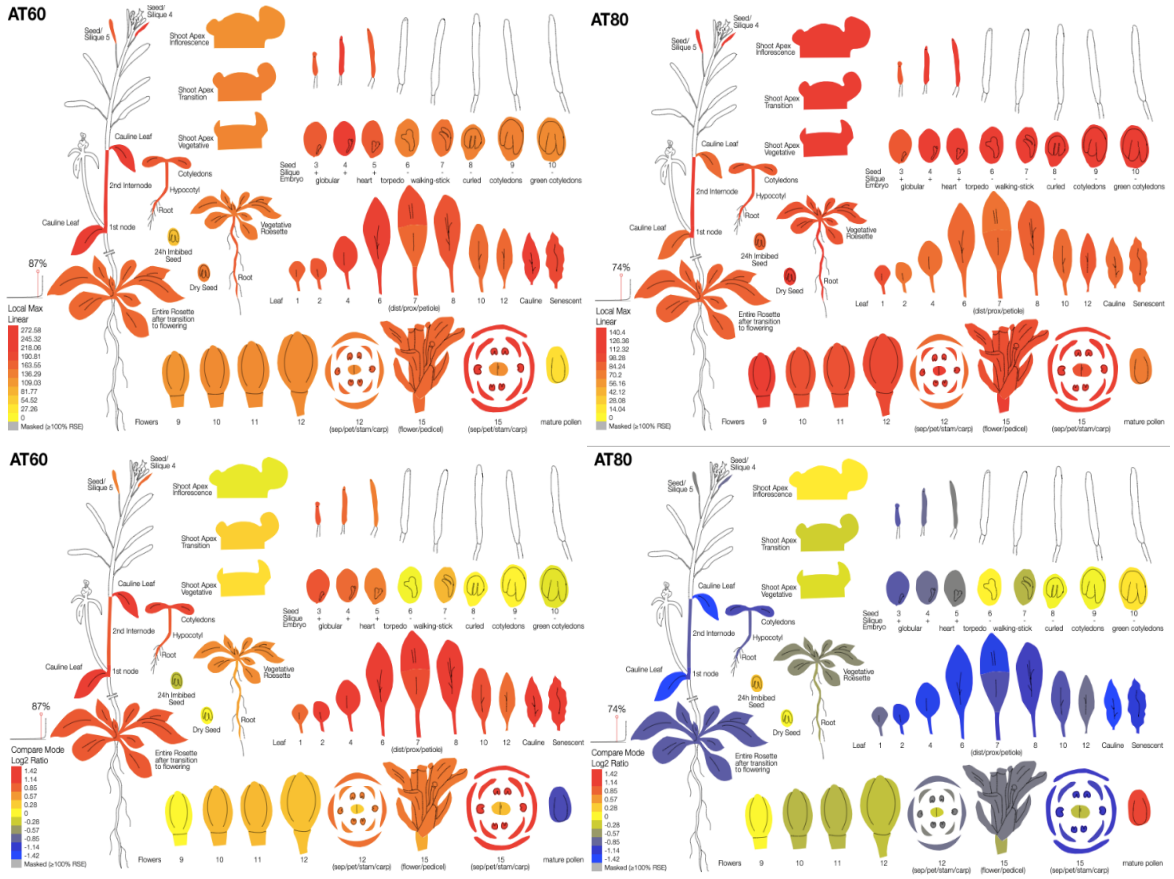


Figure 3.5 – Gene expression levels of AT60 and AT80 obtained from the AtGenExpress electronic Fluorescent Pictograph (eFP) browser. Gene expression levels and spatial patterns of AT60 (left) and AT80 (right) in normal conditions are displayed in the top panel. The comparative analysis performed on the ePlant browser are presented in the bottom panel where AT60 gene expression relative to AT80 is displayed on the left, and AT80 gene expression relative to AT60 can be seen on the right. Data originally sourced from Schmid *et al.* (2005) and Nakabayashi *et al*. (2005). Gene expression data was generated by the Affymetrix ATH1 array and normalised by the GCOS method, TGT value of 100.

3.2.2. Genetic Analyses of AT60 and AT80

Before working on the DeSi proteases for the entirety of this project, it was first required to confirm the expression of both genes in Col-0 Arabidopsis plants. To achieve this, a RT-PCR was conducted on the cDNA of Col-0 plants using specific AT80 and AT60 primers (all primer sequences are listed in Appendix Table 8.1). Firstly, RNA was extracted from Col-0 plants in their flowering stage (5-week old plants) and using this as a template, Col-0 cDNA was synthesised. Actin primers were used as a positive control. Results of this RT-PCR are shown in Figure 3.6. This experiment confirmed the expression of both proteins in Col-0 Arabidopsis plants, thereby negating any possibilities of AT60 and AT80 being a pseudogene or non-protein coding gene. As depicted in Figure 3.6, it can be deduced that the AT80 gene has greater expression levels relative to AT60.

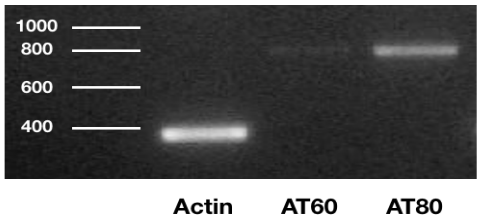


Figure 3.6 – Confirmation of AT60 and AT80 expression in Col-0 Arabidopsis Plants. RT-PCR using Taq polymerase on Col-0 cDNA with primers for actin (Actin-Fwd + Actin-Rev), AT80 (AT80-Fwd + AT80-Rev) and AT60 (AT60-Fwd + AT60-Rev).

To validate the expression levels and patterns of the AT60 and AT80 gene determined by the ePlant tool within laboratory conditions, a qPCR was conducted to quantify gene expression levels for both DeSi proteases. The actin gene was used as the housekeeping gene for calibration. This experiment was carried out across various plant structures and developmental stages of Arabidopsis plants, as well as on seedlings subjected to different stress-inducing treatments. RNA was initially extracted from Col-0 plants at different developmental stages, from various plant organs of the plant, and from 13-day old Col-0 seedlings which were transferred onto MS plates treated with ABA (25μM) and flg22 (250nM) for 8 days. These RNA samples were used to synthesise cDNA for the qPCR. This experiment would not only quantify gene expression levels for AT60 and AT80, but also investigated whether there is variation in expression levels across different plant organs and under various stress-inducing treatments. Technical repeats were conducted in triplicates, however, as the experiment was only repeated once due to time constraints, no statistical tests were conducted as a result of the limited sample size.

Results from Figure 3.7 suggests there were no discernible differences observed in the expression levels of AT60 and AT80 in 13-day old seedlings, flowers and rosette plant structures. However, it can be deduced that in the stem of Arabidopsis plants, there were considerably higher expression levels of AT80 in comparison to AT60. This can also be seen by a smaller margin in 16-day old seedlings. The opposite trend can be observed in the cauline leaves. The results except for the cauline leaf sample counter the findings from the ePlant tool, which found AT60 expression levels to be greater across all plant organs and developmental stages of Arabidopsis apart from mature pollen. Interestingly, when seedlings were exposed to ABA for 8 days, the expression of AT80 increased relative to AT80 transcript levels in 13-day old normally grown seedlings. AT80 expression levels were five times greater than AT60 levels, which experienced a 48% decrease in expression relative to normally grown seedlings. 13-day old seedlings exposed to flg22 for 8 days resulted in a decrease in both AT60 and AT80 expression levels compared to seedlings grown in normal conditions. The results from this real-time PCR experiment provide further information on the AT60 and AT80 gene expression profiles in Arabidopsis, which either aligned or contrasted with the findings from the online ePlant tool.

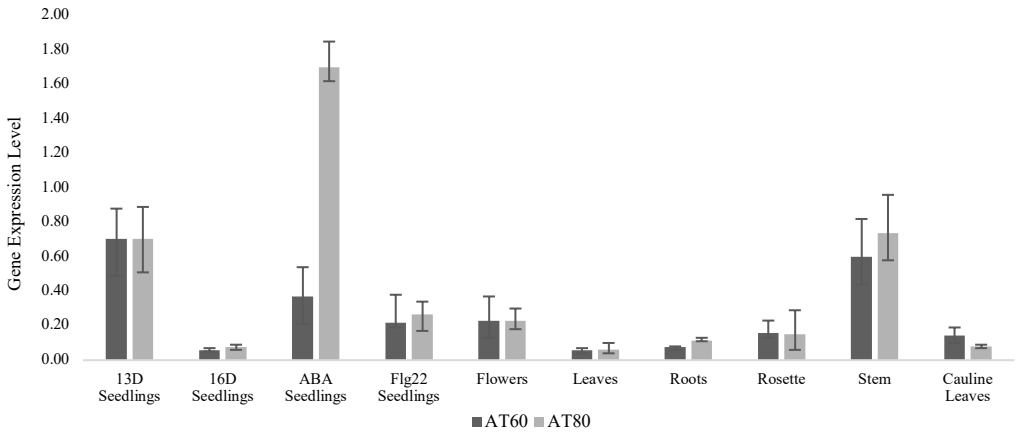


Figure 3.7 – Expression levels of AT60 and AT80 measured using real-time PCR in different plant organs and developmental stages of Col-0 plants and Col-0 seedlings subjected to various stress-inducing treatments. RNA was extracted from a range of plant organs and different developmental stages of Col-0 plants, as well as from 13-day old Col-0 seedlings grown on ABA (25μM) and flg22-(250nM) supplemented MS media plates for 8 days. cDNA was synthesised from these RNA samples and expression levels were measured using qPCR and normalised against the expression of actin in Arabidopsis. Data presented are means ± standard error (SE) of technical repeats conducted in triplicates, however, due to time constraints, the experiment was only repeated once.

3.2.3. Subcellular localisation studies of the DeSi proteases in *N. benthamiana* plants

A transient expression assay of AT60 and AT80 was conducted in *N. benthamiana* plants. This experiment was first conducted to quickly and simply determine the subcellular localisation of the DeSi proteases. *Agrobacterium* cells overexpressing the AT60 and AT80 gene within the pEG101 vector was kindly donated by Dr. Orosa. The YFP-tagged plant expression vector, pEG101, drives the transgene expression with the cauliflower mosaic virus (CaMV) 35S promoter and possesses a C-terminal YFP and hemagglutinin (HA) tag (Earley *et al.*, 2006). The *Agrobacterium* overexpressing AT60 and AT80 was syringe infiltrated into the *N. benthamiana* leaf and left for 3 days. Small 1cm^2 sections of the infiltrated leaf were cut on the 2nd and 3rd day, and the leaf samples were immediately observed under a fluorescence microscope (Leica SP5).

To prove the tag was not affecting the subcellular localisation of the proteases, an empty pEG101 vector expressing only YFP was used as a control. Total protein was extracted from each infiltrated *N. benthamiana* and analysed by SDS-PAGE and immunoblotting to detect protein presence and confirm transgene fusion to the fluorescent and epitope tags. Using antibodies raised against YFP, the immunoblot was probed. The YFP- and HA-tagged AT60 and AT80 recombinant proteins have an expected molecular mass of 58kDa and 57kDa, respectively. As seen in Figure 3.8-B, the YFP:HA:AT60 and YFP:HA:AT80 infiltrated *N. benthamiana* samples had a band this expected size confirming the respective recombinant protein was being expressed. Whereas the YFP protein band with the expected size of 27kDa was absent from the two recombinant samples, however present in the YFP control sample. This confirms the subcellular localisation seen in Figure 3.8-A is due to the YFP:AT60 and YFP:AT80 and not YFP alone.

Confocal microscopy results are displayed in Figure 3.8-A with the top panel presenting samples analysed 2 days post-infiltration and the bottom panel displaying samples analysed 3 days post-infiltration. As expected, the YFP protein localises at high intensity across both the nucleus and plasma membrane in the untransformed pEG101 vector on both day 2 and 3. Both YFP:AT80 and YFP:AT60 can be seen in abundance outside the nucleus on both days, specifically along the plasma membrane (denoted by red arrows) resembling the YFP control sample with YFP:AT60 displaying a speckled expression pattern. In both the YFP:AT60 and YFP:AT80 samples, the bright fluorescent circles (denoted by white arrows) observed in both days could be interpreted as possible protein localisation in the nucleus. However, as their appearance in comparison to the YFP control notably differs, these fluorescent circles are more likely secretory vesicles carrying the recombinant protein for exocytosis as a result of the strong protein expression driven by the CaMV35S promoter. Excluding the bright fluorescent circles, both recombinant AT80 and AT60 proteins seem to localise

faintly in the nucleus on day 2 with an appearance resembling the nuclear localisation observed in the YFP control (denoted by yellow arrows). However, by day 3, the AT60 protein is absent from the nucleus, whereas the expression of the AT80 protein is even more faint than day 2. Therefore, this analysis suggests both AT60 and AT80 localise to the cell periphery, most likely the plasma membrane, rather than the nucleus. However, this experiment lacked the necessary cellular markers to confirm the compartmentalisation of the proteases. Nevertheless, this analysis provided some indication of the approximate subcellular areas the AT60 and AT80 proteins could be found.

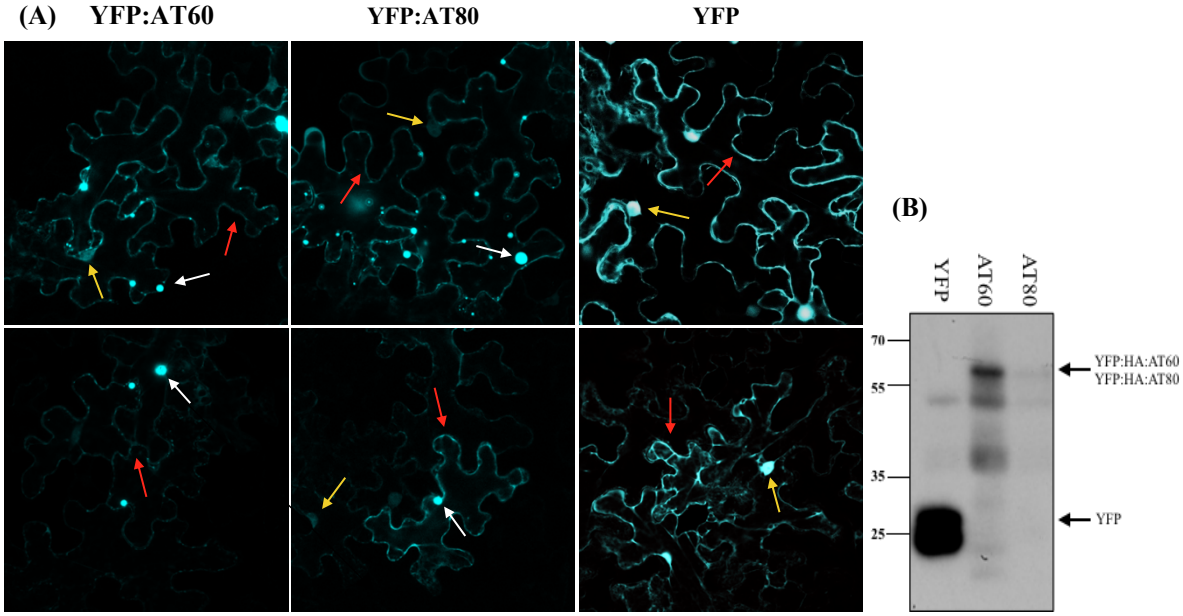


Figure 3.8 – AT60 and AT80 subcellular localisation and recombinant protein expression. A) Confocal microscopy of the cellular localisation of YFP:AT60 (left panel), YFP:AT80 (middle panel) and YFP (right panel) as the control in *N. benthamiana* plants. The top panel presents samples analysed 2 days post-infiltration and the bottom panel presents samples analysed 3 days post-infiltration. B) Total protein was extracted from infiltrated *N. benthamiana* of YFP, YFP:HA:AT60 and YFP:HA:AT80 and separated on an SDS-PAGE gel. Using anti-YFP, a western blot was performed visualise the protein bands. Coloured arrows denote subcellular components: white arrows = secretory vesicles, red arrows = plasma membrane, yellow arrows = nucleus. YFP = 27kDa, YFP:HA:AT60 = 58kDa, YFP:HA:AT80 = 57kDa.

3.3. Generation of AT60-AT80KO Transgenics with CRISPR and Mutant Analyses

As previous studies have demonstrated the pleiotropic phenotypes resulting from mutations in the SUMO conjugation or deconjugation system, this study investigates whether the AT60 and AT80 DeSi proteases also affect Arabidopsis development (Conti *et al.*, 2008). A significant proportion of the current understanding of the SUMO system was discovered through mutant analyses. To investigate the physiological functionality of the two DeSi proteases, transgenic analyses were performed. Gene KO studies are considered a major component of the functional genomics toolbox and are fundamental in revealing and understanding the role of a specific gene. In this study, the state-of-the-art targeted mutagenesis tool, CRISPR-associated protein-9 nuclease (Cas) system (Bortesi and Fischer, 2015), was utilised to perform targeted genetic KOs in Arabidopsis.

CRISPR/Cas systems originally confer molecular immunity against phages and conjugative plasmids in prokaryotes by cleaving invading foreign DNA in a sequence-dependent manner. The most commonly employed system for genetic manipulation is derived from the type II CRISPR system in *Streptococcus pyogenes* and comprises 3 components; the Cas9 RNA-guided DNA endonuclease and two identical non-coding RNA genes, trans-activating crRNA (tracrRNA) and precursor CRISPR RNA (pre-crRNA) (Bortesi and Fischer, 2015). The former provides structural stability whilst the latter consists of nuclease guided sequences (spacers), dispersed between identical direct repeats. The two RNAs bind to yield mature CRISPR RNAs (crRNAs) (Xing *et al.*, 2014). In the engineered system (Figure 3.9), the crRNA-tracrRNA complex is substituted by a single chimeric guide RNA (sgRNA) which contains a 20-nucleotide spacer sequence conferring target specificity and a hairpin structure necessary for Cas9 binding. The protospacer-adjacent motif (PAM) sequence, NGG, situated directly downstream of the target region, also confers Cas9 binding specificity. The endonuclease activity of Cas9 cleaves the DNA in the target site approximately three nucleotides upstream of the PAM site resulting in double-strand breaks (DSBs) subsequently inducing the error-prone DNA repair pathway, non-homologous end-joining (NHEJ) pathway. This repair process introduces small indels within the target region thereby resulting in mutations within the target DNA region (Wang *et al.*, 2015). As an incredibly efficient, simple and rapid genome editing tool achieving multiple gene mutations transmitting throughout the germline, CRISPR/Cas9 has successfully transformed the genomes of various organisms including bacteria, plants, animals and human cell lines (Xing *et al.*, 2014). Therefore, it was the preferred mutagenesis tool for the purposes of this project.

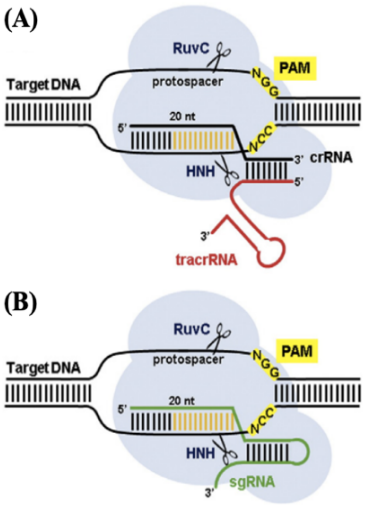


Figure 3.9 – RNA-guided endonuclease technology for mutagenesis. A) In the native system, Cas9 (blue) is guided and activated by crRNA, encompassing the 20-nucleotide spacer sequence conferring target specificity, and tracrRNA. Cas9 uses both nuclease domains (RuvC and HNH) to simultaneously cleave one strand of the target DNA at a site 3 nucleotides upstream from the PAM sequence. Of the 20-nucleotide RNA spacer sequence, the seed region (orange) is believed to promote binding between RNA and target DNA as well as confer target specificity. B) In the engineered system, sgRNA, generated by fusing the 3' end of crRNA to the 5' end of tracrRNA, can be designed to manipulate Cas9 to cleave DNA within a specified target region. The resulting DSB permits the introduction of mutations to the target DNA. Adapted from Xing *et al.* (2014).

Phylogenetic analyses revealed AT60 and AT80 share a common ancestor (Figure 3.2) and are highly homologous to one another. The pairwise sequence alignment tool revealed the DeSi proteases share high sequence similarity (Figure 3.3) and have remarkably comparable structural features as determined by their predicted 3D protein models (Figure 3.4). This suggests AT60 and AT80 most likely evolved to have similar roles and potentially function redundantly to one another. Similar to OTS1/2, it was deemed necessary that double KOs of AT60 and AT80 would need to be generated to assess the functionality of the two proteases. This section of the results chapter explains how the CRISPR protocol was performed to generate AT60 and

AT80 double KO mutant plants and following this successful transformation, the phenotypic analyses which were conducted to investigate the impact AT60 and AT80 have on Arabidopsis plant development.

3.4. Generating AT60-AT80KO Transgenic Plants with CRISPR

3.4.1. Introduction to the CRISPR/Cas9 Protocol

The CRISPR/Cas9 protocol used in this project was derived from the strategy established by Xing *et al*. (2014) and Wang *et al*. (2015). They developed a CRISPR/Cas9 vector system where binary constructs containing multiple sgRNA expression cassettes could be constructed in a single round of cloning; the Golden Gate cloning method. The cloning vector, pCBC-DT1T2 created by Xing *et al*. (2014), was used as a template for producing expression cassettes with multiple sgRNA target sites. Wang *et al*. (2015) developed a recombinant destination vector, pHEE401E, based on the pCAMBIA backbone, which consists of an egg cell-specific promoter, *EC1.2*, to drive the Cas9 expression. This CRISPR/Cas9 binary vector set has successfully generated T1 homozygous mutants in Arabidopsis and the mutations were germline inherited by the following generation. Therefore, this vector system was selected for this study.

The simplest and most efficient way in which to perform a double KO experiment within a limited time frame was to utilise the CRISPR/Cas9 system to KO one gene within a stable, homozygous transfer DNA (T-DNA) mutant of the other gene. TAIR database was used to identify a homozygous mutant which would be used as the background genome for the CRISPR KO experiment. Fortunately, a homozygous mutant of the AT80 gene (SALK_064598C) had been isolated and donated by Joseph R. Ecker from The Salk Institute for Biological Studies (Alonso *et al.*, 2003). Furthermore, as the selected CRISPR/Cas9 binary vector set could deliver two sgRNAs for targeted genetic mutations, this project experimented whether dual targeted genetic mutations could be achieved using the CRISPR/Cas9 binary system. An expression cassette with two sgRNAs targeting the AT60 and AT80 gene was assembled and transformed into Col-0 plants. This was conducted as a peripheral experiment to test whether the selected vector system could achieve the generation of AT60-AT80KO mutant plants. The CRISPR/Cas9 protocol by Xing *et al.* (2014) and Wang *et al.* (2015) also greatly suited the purpose of this experiment, as it accelerated the application of CRISPR/Cas9 for targeted genetic mutations in the background of WT and T-DNA mutant Arabidopsis plants. Furthermore, the dual sgRNA approach where a gene is targeted by two sgRNAs simultaneously results in more reliable loss-of-function alleles and therefore was the optimum method for generating KO mutants (Pauwels *et al.*, 2018). For delivering the plasmids containing Cas9 and multiple sgRNA expression cassettes into Arabidopsis plants, the *Agrobacterium*-mediated transformation method was performed. DNA sequencing was conducted to analyse the T1 transgenic plants for homozygous mutations.

3.4.2. Generating Expression Cassettes for sgRNAs

Two expression cassettes were constructed for this project. The first expression cassette comprised of two sgRNAs targeting the AT60 gene that would be transformed into the AT80 T-DNA homozygous mutant to generate an AT60-AT80KO mutant plant. The second expression cassette, which was conducted as a peripheral experiment, comprised of two sgRNAs with one targeting the AT60 gene and the other targeting the AT80

gene. This cassette would be transformed into Col-0 plants to determine whether the dual sgRNA expression cassette is capable of mutating two genes simultaneously, resulting in an AT60-AT80 double KO mutant.

3.4.2.1. Identification of sgRNA Spacer Sequences

Identifying a spacer sequence for each sgRNA is crucial for specific target DNA binding and cleavage, as well as for preventing off-target binding events (Jiang and Doudna, 2017). The Portal of CRISPR-Cas9 Mediated Genome Editing (https://www.genome.arizona.edu/crispr/CRISPRsearch.html) was used to identify the 20-nucleotide spacer sequence directly upstream of the Cas9 PAM site (5'-N_{20}NGG-3') for the AT60 and AT80 gene within the Arabidopsis genome. The spacer sequences were searched throughout the exon regions of the genomic DNA sequence for each gene. The selected spacer sequences are listed in Table 3.1.

Table 3.1 – Spacer sequences and PAM sites for sgRNAs targeting AT60 and AT80. The sequences highlighted were used in the subsequent protocol step. The PAM sites for the sgRNAs are in bold and in red.

Spacer	Gene	
	AT60	AT80
Spacer 1	5' – ACGTCTGAGATTGAAAAGTT**TGG** – 3' 5' – **CCA**AACTTTTCAATCTCAGACGT – 3'	5' – TACGTCTGAAATTGAGAAAT**CGG** – 3' 5' – **CCG**ATTTCTCAATTTCAGACGTA – 3'
Spacer 2	5' – GCAGAAGTGATTGCAATTTT**TGG** – 3' 5' – **CCA**AAAATTGCAATCACTTCTGC – 3'	5' – TGCCAAGGACGATTTTCTCA**CGG** – 3' 5' – **CCG**TGAGAAAATCGTCCTTGGCA – 3'

For both the AT60 and AT80 gene, the first spacer sequences were located on the positive strand of an exon, whilst the second spacer sequences were situated on the negative strand. As the second spacer is reversely inserted into the expression cassette, the reverse complementary sequence was selected. For the dual sgRNA expression cassette, the first spacer sequence for AT60 and the second spacer sequence for AT80 were selected. The selected spacer sequences for the sgRNAs were evaluated for their target specificity using the Cas-OFFinder website which searches for potential off-target sites of Cas9 RNA-guided endonucleases (http://www.rgenome.net/cas-offinder/).

To assemble the expression cassettes with the selected spacer sequences, 19-nucleotides of the first spacer sequence directly adjacent to the PAM site was incorporated into two PCR forward primer templates (BsF and F0). The second spacer sequence (reverse-complemented) directly adjacent to the PAM site was incorporated into two PCR reverse primer templates (BsR and R0). The primers used are listed in Appendix Table 8.1.

3.4.2.2. Assembly of Expression Cassettes by PCR

The pCBC-DT1T2 plasmid was used as a template to assemble the sgRNA expression cassette. The PCR reaction was adapted from the protocol by Xing *et al*. (2014) and the reactions were amplified using the Q5® High-Fidelity DNA Polymerase (Table 3.2). The pCBC-DT1T2 plasmid was kindly donated by Dr. de Lucas and the plasmid map can be found in Appendix Figure 8.1.

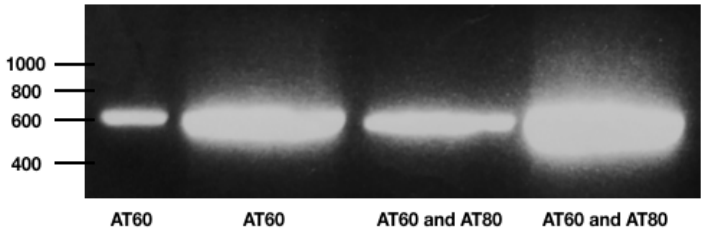
Table 3.2 – PCR reaction mix and reaction conditions to assemble sgRNA expression cassettes. The N represents the gene name for the forward primers (BsF and F0) containing the first spacer sequence, and the reverse primers (BsR and R0) encompassing the second spacer sequence.

Components	Volume	Reaction Conditions
5x Q5® Reaction Buffer (NEB)	10µl	98°C – 40 seconds
dNTPs (2mM)	4µl	

Q5® High-Fidelity DNA polymerase (NEB)	1µl	98°C – 15 seconds 70°C – 30 seconds 72°C – 20 seconds	x35 cycles
pCBC-DT1T2 plasmid (2ng/µl)	1µl		
N-BsF (20uM)	1µl		
N-F0 (1uM)	1µl		
N-R0 (1uM)	1µl		
N-BsR (20uM)	1µl		
Sterile Distilled Water	30µl	72°C – 2 minutes	

The first PCR reaction was to assemble the sgRNA expression cassette containing two sgRNAs targeting the AT60 gene. Therefore, the reaction consisted of the primers: AT60-BsF, AT60-F0, AT60-BsR and AT60-R0. The second PCR reaction was to make the sgRNA expression cassette with two sgRNAs dually targeting the AT60 and AT80 gene. Therefore, the reaction mixed the forward primers targeting AT60 (AT60-BsF and AT60-F0) and the reverse primers targeting AT80 (AT80-BsR and AT80-R0). The results of the PCR reactions were visualised using gel electrophoresis and the expected gel fragment size was 624 base pairs (bp) (Figure 3.10). The PCR product was then extracted from the gel and purified so it could be used in the next step.

Figure 3.10 – Gel electrophoresis for PCR products containing sgRNAs targeting AT60 and AT60-AT80. Each reaction yielded the expected 624bp product by using pCBC-DT1DT2 as a template and specific primers. The AT60-targeting sgRNA was assembled using the primers: AT60-BsF, AT60-R0, AT60-BsR and AT60-R0. The AT60- and AT80-targeting sgRNA was generated using the primers: AT60-BsF, AT60-F0, AT80-BsR and AT80-R0.



3.4.3. Golden Gate Reaction and Transformation of Recombinant Plasmids

To digest and ligate the purified PCR product comprising the dual sgRNAs targeting AT60 alone or AT60 and AT80, the Golden Gate reaction was performed. The extracted PCR product was inserted between the BsaI sites of the pCAMBIA plasmid, pHEE401E, which was kindly donated by Dr. de-Lucas and the plasmid map can be found in Appendix Figure 8.1. The reaction protocol listed in the methods section was based on Xing *et al*. (2014) paper with minor modifications. This golden gate reaction was set up to assemble the sgRNA expression cassettes targeting AT60 and AT60-AT80.

5µl of the golden gate reaction mix was then transformed into DH5α *E. coli* cells. This was conducted for both sgRNA expression cassettes (targeting AT60 and AT60-AT80). A colony PCR was conducted on the colonies growing on the kanamycin (50µg ml^{-1}) LB agar selection plates following transformation (Appendix Figure 8.2 and 8.3). Positive colonies were then selected to make overnight cultures for plasmid isolation using the miniprep protocol. The isolated recombinant plasmids were then sequenced using the M13-Fwd primer to confirm the spacer sequences for each target gene were inserted correctly into the plasmids (Appendix Figure 8.4).

1µg of the recombinant plasmids containing the sgRNAs were transformed into competent *Agrobacterium* cells. This was conducted for both sgRNA expression cassettes (targeting AT60 and AT60-AT80). A colony PCR was conducted on the colonies from the gentamicin (10µg ml^{-1}), kanamycin (50µg ml^{-1}) and rifampicin

(10µg ml^{-1}) LB agar plates following transformation (Appendix Figure 8.5). Positive colonies were grown in overnight cultures and used to make glycerol stocks stored at -80°C.

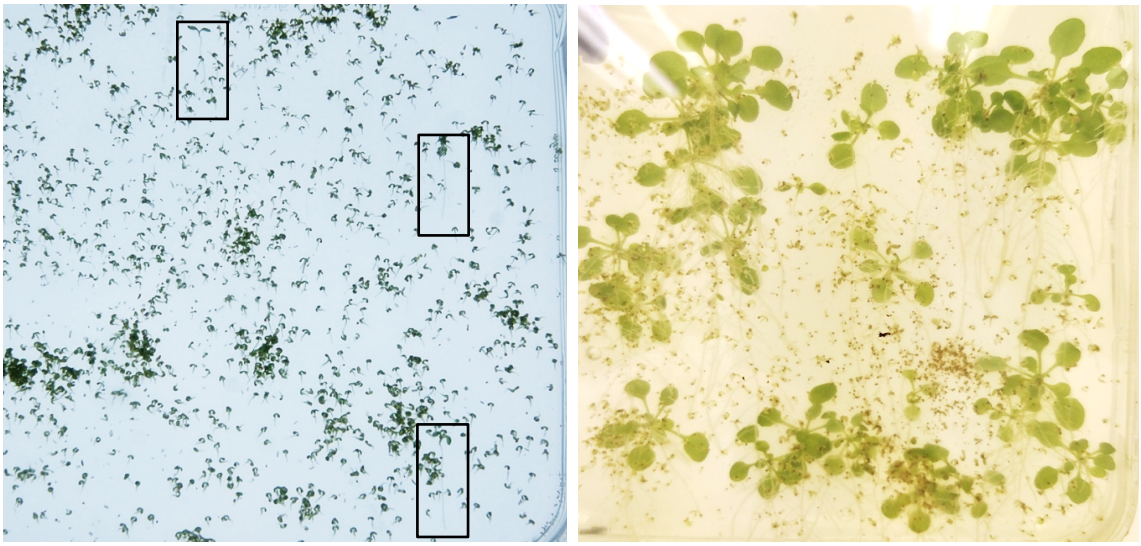
3.4.4. *Agrobacterium*-mediated Transformation of Arabidopsis by Floral Dipping

The glycerol stocks of successfully transformed recombinant *Agrobacterium* cells carrying the sgRNA expression cassettes were grown in 10ml overnight cultures with gentamicin (10µg ml^{-1}), kanamycin (50µg ml^{-1}) and rifampicin (10µg ml^{-1}). Subsequently, the cultures were scaled up and grown in 1L of LB liquid. The floral dipping method for *Agrobacterium*-mediated transformation of Arabidopsis was conducted by following the protocol stated in the methods. Two-month-old flowering Arabidopsis plants with cut siliques were then transformed by the recombinant *Agrobacterium*-medium prepared. The homozygous AT80KO mutant plant was transformed with *Agrobacterium* containing sgRNAs targeting the AT60 gene. Col-0 plants were transformed with *Agrobacterium* containing sgRNAs dually targeting the AT60 and AT80 gene. Following floral dipping, the plants were grown in normal conditions and all seeds were subsequently harvested.

3.4.5. Screening for T1 Transgenic Plants

Seeds harvested from the T0 plant were sterilised and sown on MS plates containing 30µg ml^{-1} of hygromycin. After approximately two weeks, plates were screened for transformants. Resistant seedlings were then transferred to soil and grown in normal conditions (Figure 3.11). At this stage, there were 9 potential AT60KO in AT80KO background transformants and 4 potential AT60-AT80KO in Col-0 background transformant plants. Once each of the potential transformant lines reached the rosette stage and genomic DNA could be extracted from the leaves, mutant genotyping analyses were conducted.

(A) AT60KO in AT80KO Background **(B)** AT60-AT80KO in Col-0 Background



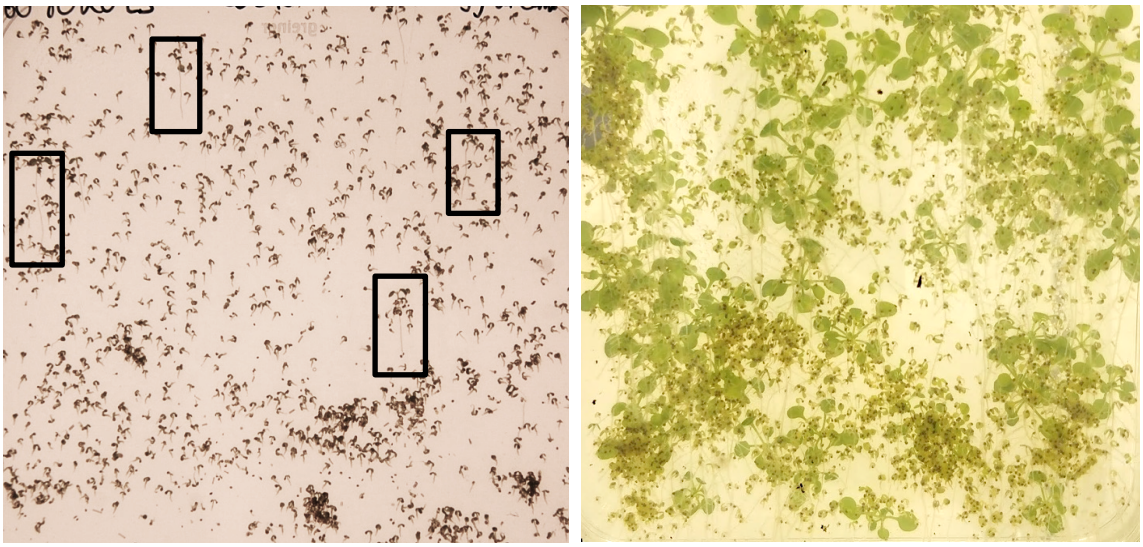


Figure 3.11 – Screening two-week old T1 seedlings for transformants. Seeds were harvested from *Agrobacterium*-transformed T0 plants and sown on hygromycin (30µg ml^{-1}) MS selection plates. Resistant seedlings were transferred to soil and grown. A) T1 seedlings transformed with sgRNA targeting AT60 in AT80KO background. The black boxes outline the resistant seedlings. B) T1 seedlings transformed with sgRNA dually targeting AT60 and AT80 in Col-0 background. The seedlings with larger leaves and longer branching roots were considered transformants.

3.5. Genotyping Analyses Conducted on KO Mutant Transgenic Plants

3.5.1. PCR Confirming the Homozygous AT80 T-DNA KO Mutant

The AT80KO mutant plants were genotyped by PCR to confirm the T-DNA insertion was present before the *Agrobacterium*-mediated transformation. Genomic DNA was extracted from flowering AT80KO mutant plants. To identify the T-DNA insertion deleting the AT80 gene, a T-DNA left border-specific primer was used with an AT80 gene specific forward primer. To check the mutants were homozygous KOs for the AT80 gene, specific primers isolating the AT80 gene were also used. PCR reactions isolating the AT80 gene and T-DNA with AT80 were conducted on the genomic DNA extracted from three AT80KO T-DNA mutant lines (AT80KO #1, #2 and #3) and Col-0 plants. PCR products were run on gel electrophoresis and results are displayed in Figure 3.12. For Col-0, the expected band size for the fragment isolating the AT80 gene (784bp) can be observed with no band present for the T-DNA and AT80 gene fragment. For AT80KO mutants #1 and #2, a fragment flanking the T-DNA border and AT80 gene (approximately 1300bp) can be seen, whilst no band was present for the AT80 gene, confirming the mutants were carrying T-DNA insertions knocking out the AT80 gene. The AT80KO #3 mutant had a band corresponding to the uninterrupted strand of the AT80 gene in addition to the T-DNA disrupted strand of the AT80 gene indicating this line was a heterozygous AT80KO mutant. Therefore, the genotyped mutants confirmed as homozygous AT80KO lines (AT80KO #1 and #2), were used to generate AT60-AT80KO mutants.

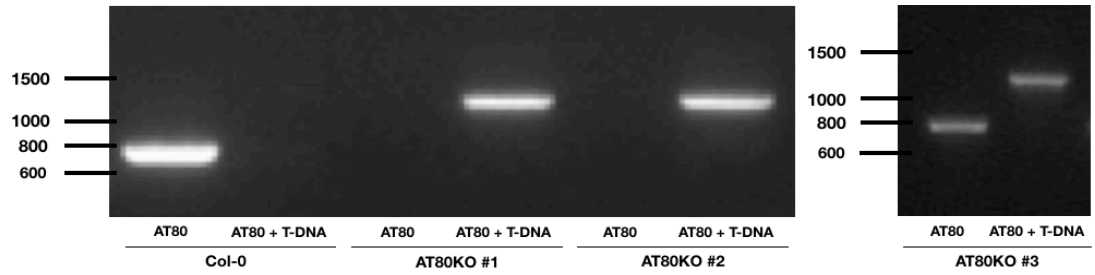


Figure 3.12 – PCR results confirming T-DNA insertion in the homozygous AT80KO mutant plants (SALK_064598C). Three T-DNA KO mutants of AT80 (AT80KO #1, #2 and #3) and Col-0 plants were genotyped by PCR using AT80 specific primers (AT80-Fwd + AT80-Rev) and a T-DNA left border-specific primer (SALK T-DNA LB) with the AT80 forward primer (AT80-Fwd). The expected band size for the AT80 gene fragment was 784bp. The band size of the PCR product flanking the T-DNA insertion of the AT80 gene was approximately 1300bp.

3.5.2. Sequencing Results of the T1 Transgenic KO Plants

To confirm the T1 transgenic plants were KO mutants, RT-PCR analysis was initially performed. The expected results from the RT-PCR would show an absence of the full-length AT60 and AT80 mRNA in the T1 transgenic plants. This RT-PCR analysis used positive controls such as actin primers to isolate the actin gene and Col-0 samples. The results are shown in Appendix Figure 8.6. However, this experiment was regarded inconclusive as the assay depended on a negative result to confirm the KO mutations, as well as the fact that CRISPR-Cas9 only produces a few deletions or insertions in the target DNA sequence of the gene and therefore may still translate the targeted gene into mRNA. To conclusively determine whether the target genes have been successfully knocked-out in the T1 transgenic plants, the region flanking the target regions were sequenced. This analysis would also enable the CRISPR-Cas9 mutations to be identified.

For sequencing, the genomic DNA was extracted from the leaves of the rosette stage T1 transgenic plants for both AT60KO in AT80KO background and AT60-AT80KO in Col-0 background. Specific primers were designed to isolate the target region flanking the sgRNAs. AT60-G1F and AT60G2F primers isolated the first sgRNA region targeting the AT60 gene for both transgenic plant lines. For the second sgRNA region targeting the AT80 gene for the AT60-AT80KO in Col-0 background lines, AT80-G1F and AT80-G1R primers were used. Once the sequencing results were obtained, the sequences were initially analysed using the SnapGene Viewer software to manually double-check the interpretation of the primary data and correct any mis-called nucleotides. Subsequently, the sequences for the T1 transgenic lines were aligned with the WT sequence of the corresponding target region using Clustal W to identify indel mutations in the transgenic plant sequence.

All 9 T1 potential AT60KO in AT80KO transformants and all 4 T1 potential AT60-AT80KO in Col-0 transformants were sequenced. The sequence alignment results for three lines of the T1 AT60KO in AT80KO mutants and one line of the T1 AT60-AT80KO in Col-0 are displayed in Table 3.3. The CRISPR/Cas9 system usually generates mutations close to the DSB site which occurs three base pairs upstream of the PAM site. As can be seen in Table 3.3, insertions, deletions and substitutions have taken place in the expected targeted DNA region located near the DSB site (highlighted in yellow) in all T1 transgenic lines, with some lines displaying mutations closer to the DSB site than others. For instance, the T1 transgenic line, 60KO#1 (A1) has an insertion

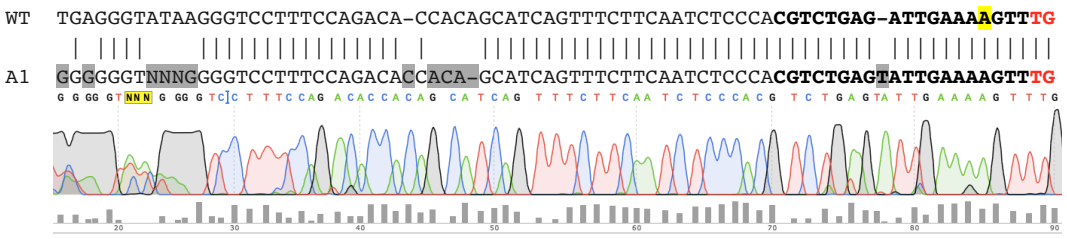
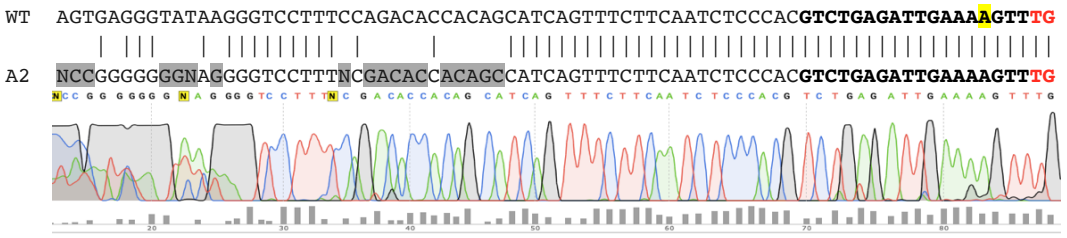
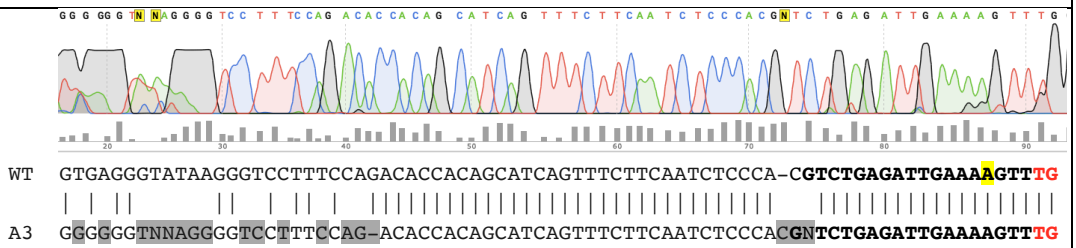
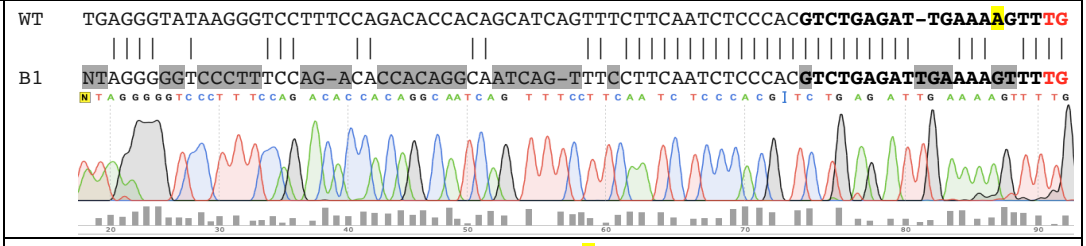
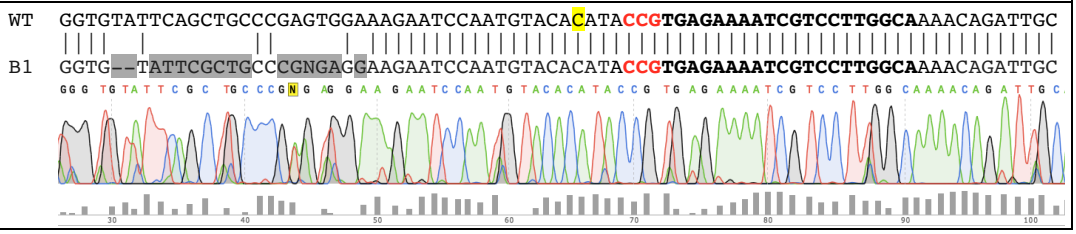
8 nucleotides upstream from the DSB site whilst the 60-80KO (B1) line has an insertion 6 nucleotides upstream from the DSB site. However, as genomic sequencing is prone to misalignment and miscalls as a consequence of sample quality and the sequencing analysis was not biologically or technically repeated, the mutations observed in the T1 transgenic lines could not be conclusively attributable to the CRISPR-Cas9 system. Furthermore, the differences identified in the T1 transgenic lines from WT further upstream from the target site could be due to the low-quality start of the electropherogram rather than the CRISPR-Cas9 system. If more time persisted, multiple biological and technical repeats would be conducted, as well as the generation of more KO transgenic lines for the screening process. Therefore, although mutations were identified within the sequenced T1 transgenic lines, at this stage due to limited evidence, they could only be presumed to be CRISPR-Cas9 KO mutants and the possibility of unsuccessful transformation should be considered in later experimental steps. However, due to the given time constraints the CRISPR-Cas9 protocol and sequencing analysis could not be repeated, and the T1 transgenic lines were taken forward to subsequent generations for further analysis.

As CRISPR generates indel mutations which can be homozygous or heterozygous, the pattern of zygosity was identified by analysing the sequence chromatogram data in the SnapGene Viewer software. Within a single peak position, if a double peak representing two different nucleotides is present, it can be inferred as a heterozygous indel or single nucleotide polymorphism (SNP). The peak positions of the indel mutations identified in the T1 transgenic line sequences were examined to determine whether they were heterozygous or homozygous indels. All indel mutation positions for each of the T1 transgenic lines only contained a single peak, therefore denoting the indel mutations generated by CRISPR-Cas9 were homozygous. These results are presented below the sequence alignments for each of the transgenic lines (Table 3.3). Although this method was useful in providing information on the sequencing data and an indication on the type of indel mutation that occurred in the T1 transgenic lines, it was completely dependent on the sample quality and prone to human error. Therefore, this method was not conclusive in determining the zygosity of the CRISPR-Cas9 T1 mutants, however, it was the best alternative within the given time constraints.

From this experiment, it could only be inferred the CRISPR-Cas9 system was successful in generating mutations within the target genes in all tested T1 transgenic lines and that dual sgRNA targeting could mutate two genes simultaneously. However, the possibility of unsuccessful transformation should be considered in later experimental stages. Analysing the DNA sequencing chromatograph data of the T1 transgenic lines indicated the indel mutations generated by CRISPR-Cas9 were homozygous for all lines, however this was again inconclusive. The T1 mutant lines listed in Table 3.3 were subsequently taken to the T2 and T3 generation for phenotyping purposes as well as to further determine the zygosity of the mutant lines.

Table 3.3 – Sequence alignment of T1 transgenic lines with WT by Clustal W and DNA sequencing chromatogram data. Mutations in the transgenic lines are highlighted in grey and the spacer sequences are in bold. PAM sites are in red and the expected DSB sites are highlighted in yellow. AT60KO denotes the transgenic lines with CRISPR KOs in the AT60 gene in AT80KO background. Results only show the sequence of the first spacer sequence of the sgRNA targeting the AT60 gene aligned with the WT sequence. AT60-80KO represents the transgenic line with KOs in the AT60 and AT80 gene in Col-0 background. Results show the sequence for the first and second spacer sequence of the sgRNA targeting AT60 (top) and AT80 (bottom)

respectively, aligned with the WT sequence. Below the sequence alignments, the chromatograph data for the T1 transgenic line is displayed to facilitate identifying whether the mutations are homozygous or heterozygous, along with the estimated quality of the data below.

T1 Line	Sequence Alignment and Chromatogram Data
60KO#1 A1	
60KO#2 A2	
60KO#3 A3	
60-80KO B1	
	

3.6. Phenotypic Characterisation of AT60-AT80KO Mutant Lines

3.6.1. Screening T2 and T3 Transgenic Plants

The sequenced T1 transgenic lines inferred as homozygous AT60-AT80KO mutants were propagated to T2 generation. This was done by harvesting seeds from the T1 transgenic lines and screening them on hygromycin (30µg ml^{-1}) MS selection plates. The T2 resistant seedlings were subsequently subjected to phenotyping experiments including root length and fresh weight assays, as well as grown in normal conditions to be taken

to T3. Once again, the seeds harvested from the T2 transgenic lines were screened on hygromycin (30µg ml^{-1}) MS selection plates. This time they were screened for 100% survival. The T3 lines with 100% resistant seedlings could be considered to be homozygous KO mutants. T3 transgenic seedlings from lines segregating at 100% were then subsequently subjected to phenotyping experiments. The results from this segregation analysis are displayed in Table 3.4. The T3 transgenic lines, A2-8, A2-3, A1-3 and A1-6 were homozygous mutants. Therefore, the T3 A2-8 and A1-3 lines were selected for a second root length assay.

Table 3.4 – Segregation analysis results from screening T3 transgenic lines on hygromycin MS selection plates. Seedlings harvested from individual T2 transgenic lines were grown on hygromycin (30µg ml^{-1}) MS selection plates for 6 days. The approximate percentage of resistant seedlings were calculated by the ratio of resistant to non-resistant seedlings. Zygosity of the T3 transgenic seedlings could be inferred by segregation percentage. Images of the screened selection plates can be observed in Appendix Figure 8.7.

T3 Transgenic Line	Approximate Percentage of Resistant Seedlings	Zygosity
A1 – 3	100%	Homozygous
A1 – 5	75%	Heterozygous
A1 – 6	100%	Homozygous
A2 – 3	100%	Homozygous
A2 – 8	100%	Homozygous
A2 – 10	75%	Heterozygous
A3 – 1	75%	Heterozygous
A3 – 6	75%	Heterozygous
B1 – 1	75%	Heterozygous
B1 – 3	75%	Heterozygous
B1 – 4	75%	Heterozygous

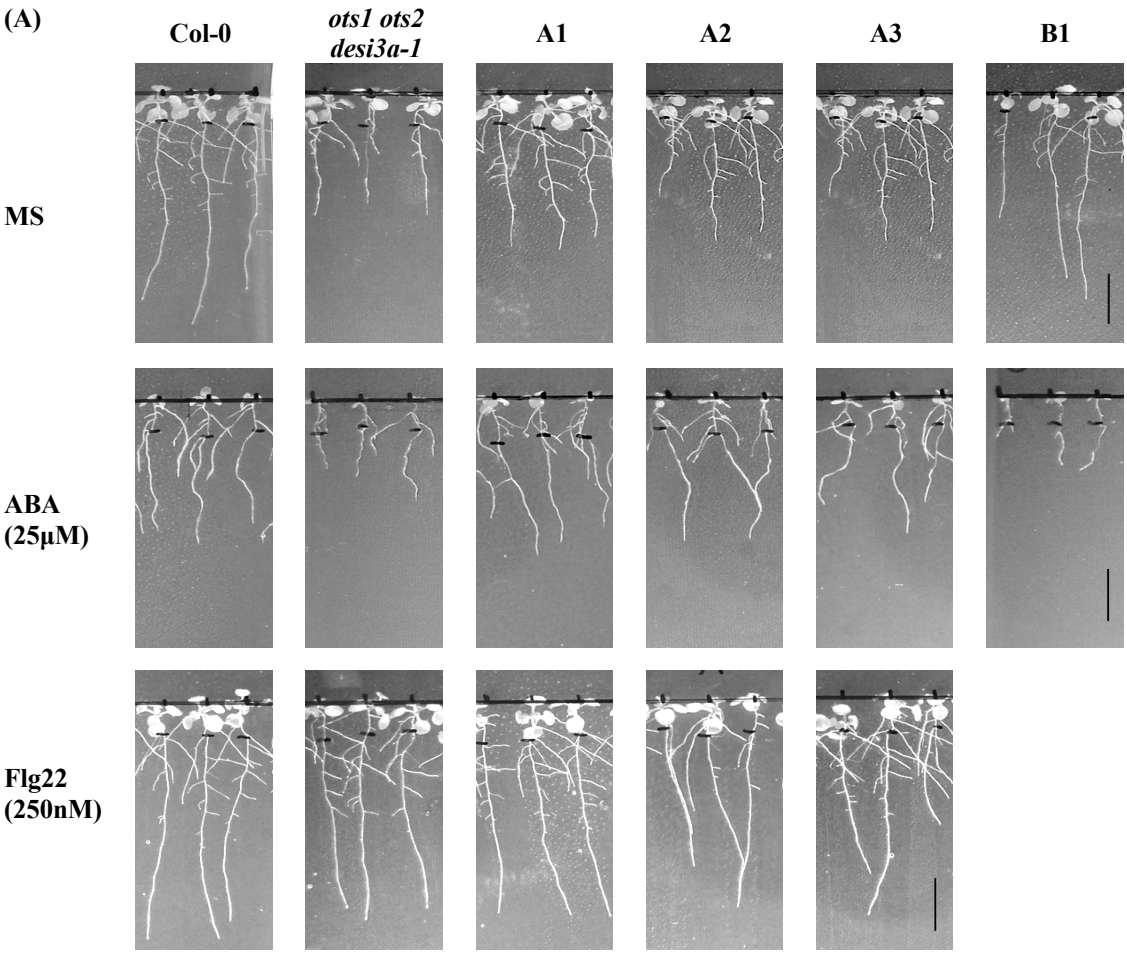
3.6.2. Root Length and Fresh Weight Assays on T2 Mutant Transgenic Plants

To phenotype the mutant plants, a root length and fresh weight assay in response to normal and stress conditions were performed. These assays are one of the most effective and time-efficient experiments used to identify and investigate the role of the genes of interest in plant development. This section assesses whether the T2 AT60-AT80KO mutants display pleiotropic phenotypes in root growth and biomass under normal conditions and in response to the stress inducers, ABA and pathogen response elicitor, flg22.

3.6.2.1. Root Length Assay on T2 KO Transgenic Lines

Col-0, *ots1 ots2*, *desi3a-1* and AT60-AT80KO mutant lines were tested in the root growth and fresh weight assay. The *ots1 ots2* and *desi3a-1* lines were used as comparative samples as both were KO mutants of Arabidopsis SUMO proteases, particularly the latter being a DeSi SUMO protease. The seedlings for each line were first grown on normal MS media for 4 days and then transferred to MS media, ABA (25µM) containing MS media and flg22 (250nm) containing MS media. As a recent study established a role for Desi3a in flagellin sensing (Orosa *et al.*, 2018), *desi3a-1* was tested on flg22-supplemented media whereas *ots1 ots2* was tested on ABA-supplemented media. 6 days following the transfer, root lengths of each seedling for each genotype were measured digitally using the Fiji software. Representative samples for each genotype on normal MS, ABA MS and flg22 MS media are displayed in Figure 3.13-A. Quantification of the root lengths were calculated using averages of all seedlings analysed per genotype. At least 10 individual seedlings for each genotype were tested, however, the experiment was only repeated once due to time constraints. The average root length for each genotype grown on all three mediums are presented in Figure 3.13-B. The Kruskal-Wallis

test followed by a Bonferroni post hoc test was used to calculate statistical significance. Results are displayed in Figure 3.13-B. To quantify the effect ABA and flg22 had on the root elongation for all lines, the difference between the average root length of seedlings grown on MS media and seedlings grown on ABA or flg22-supplemented media of the same genotype was calculated. Results of this analysis are displayed in Figure 3.13-C and statistical significance was calculated using the Mann-Whitney U-test which compared the root lengths between normally grown and treatment-exposed seedlings. This analysis determined whether the difference in root length was caused by overall slower growth of the genotype or the presence of ABA or flg22 restricting root elongation of the seedlings.



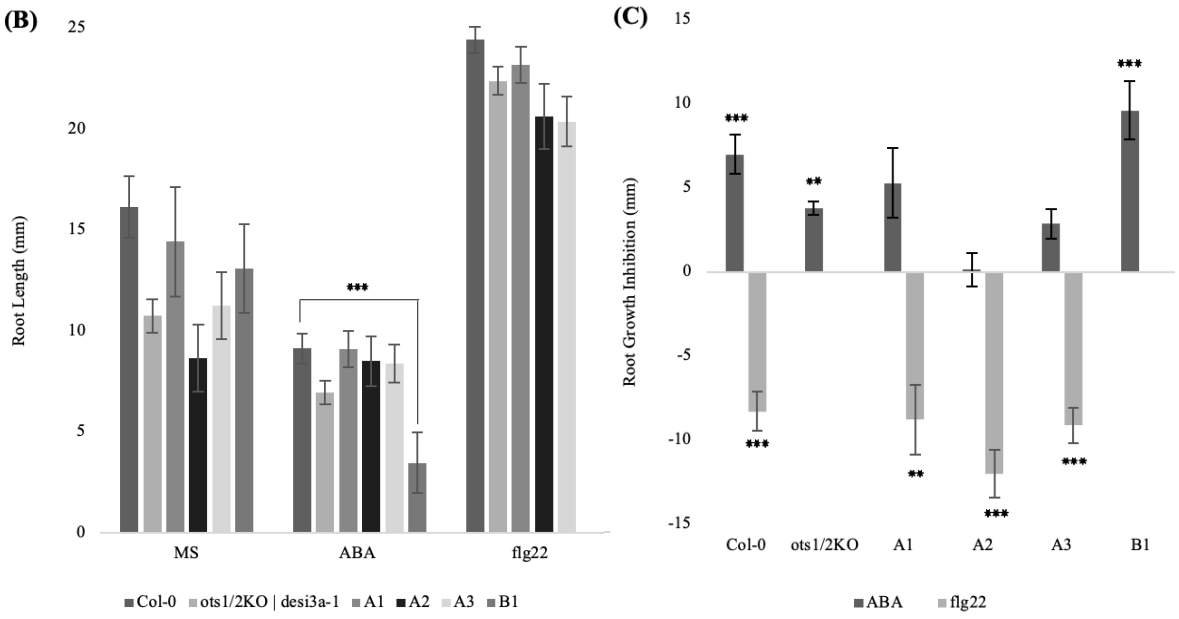


Figure 3.13 – Results from the T2 Root Length Assay. A) Phenotypic appearance of representative 10-day old seedlings from each genotype grown on MS media, and ABA (25µM) and flg22 (250nm) supplemented MS plates for 6 days. Scale bar equals 1cm. B) Quantification of average root growth for each genotype grown on normal and treatment-containing media. C) Quantification of average root growth inhibition for each genotype comparing the difference between root lengths of seedlings grown in normal and stress-induced conditions (ABA and flg22). Data presented are means ± standard error (SE) from at least 10 individual seedlings for each genotype. Significance was assessed using Mann-Whitney U test for 2 samples (C) and Kruskal-Wallis test followed by Bonferroni post-test for comparing 3 samples or more (B). Significance values: * = $P < 0.05$; ** = $P < 0.01$; *** = $P < 0.005$.

When comparing between genotypes grown on each medium, the Kruskal-Wallis test determined the mean root length differed significantly between genotypes that were grown on ABA-treated media. Post hoc tests using the Bonferroni correction revealed that Col-0 grew significantly longer than B1 when subjected to ABA conditions. Aside from this, all lines displayed similar root elongation under normal and stress-induced conditions. This confirms the T2 transgenic lines under normal and flg22 conditions do not significantly differ in root length from WT. The Mann-Whitney U test for the root growth inhibition assay found that ABA significantly repressed the root elongation of Col-0, *ots1 ots2* and B1 lines. However, ABA had not significantly affected the root elongation of the other three T2 KO mutant lines. On the other hand, flg22 treatment significantly promoted root elongation for the three T2 mutant lines as well as for Col-0. These results suggest that AT60-AT80KO mutant plants could be more insensitive to ABA than WT.

3.6.2.2. Fresh Weight Assay on T2 KO Transgenic Lines

To further elucidate the phenotypic difference between the genotypes in normal and stress-induced conditions, fresh weight data was collected from 10-day old seedlings. This assay establishes whether the biomass of the AT60-AT80KO mutant varies from WT and other SUMO protease KO mutants (*ots1 ots2* and *desi3a-1*) in normal and stressed conditions. It also discerns whether the phenotypes observed in the root length assay was a consequence of the stress inducer triggering the switch from growth to defence mode inhibiting root growth

mechanisms or global growth of all plant tissues. The seedlings grown on MS and ABA- and flg22-supplemented MS media for 6 days were weighed and quantified using the averages of all analysed seedlings per genotype. The graphs were split and displayed by their comparative groups in Figure 3.14.

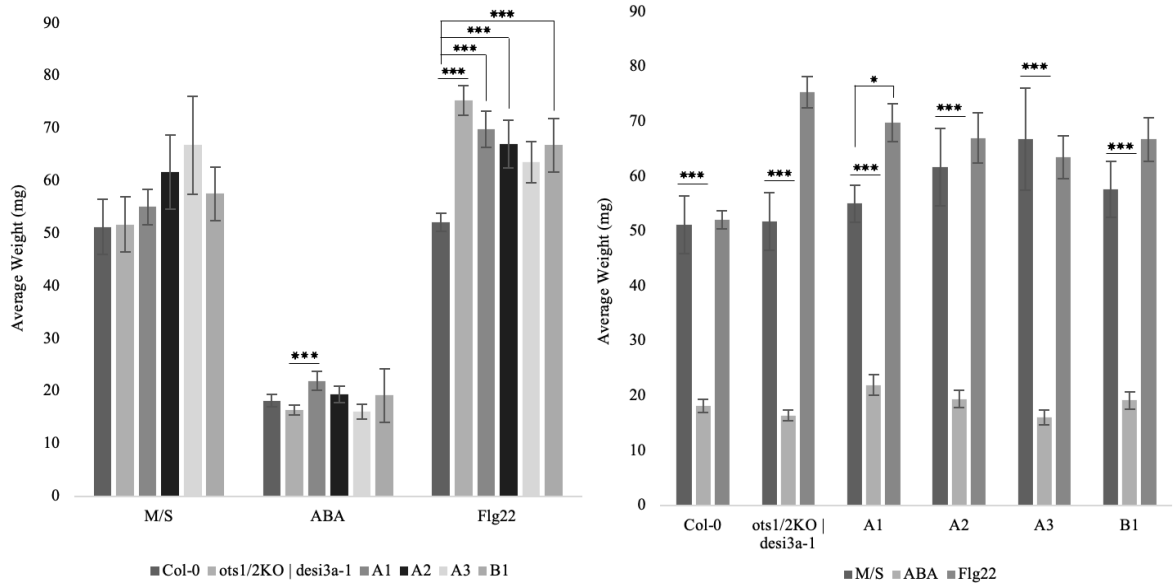
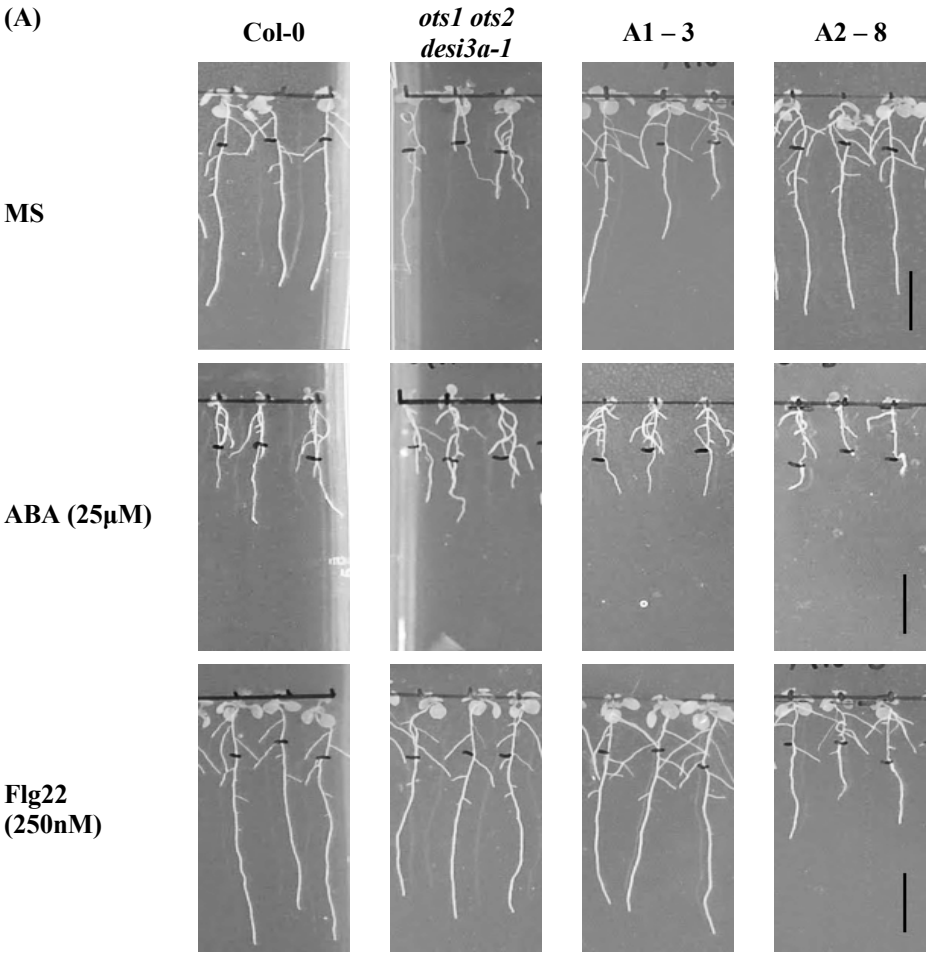


Figure 3.14 – Results from the Fresh Weight Assay. Quantification of the biomass of 10-day old seedlings exposed to treatment for 6 days. Seedlings were grown on MS media for 4 days then transferred to MS plates with ABA (25µM) and flg22 (250nM) and without. After 6 days of further growth, the plants were weighed. The left graph compares the average fresh weight of seedlings between genotypes for each medium. The right graph compares the effect of ABA and flg22 on each genotype. Data presented are means ± SE from at least 10 individual seedlings for each genotype. Significance was assessed using the Mann-Whitney U test for 2 samples and Kruskal-Wallis test followed by Bonferroni post-test for comparing 3 samples or more. Significance values: * = $P < 0.05$; ** = $P < 0.01$; *** = $P < 0.005$.

The Kruskal-Wallis test determined no statistically significant difference in average biomass between the genotypes of seedlings in normal conditions, thereby denoting there were no biomass differences between WT and the AT60-AT80KO mutants. The statistical test did find significant differences between genotypes in the biomass of seedlings grown on ABA. Bonferroni post-hoc test revealed A1 was significantly heavier than *ots1 ots2* when exposed to ABA. The fresh weight of seedlings varied dramatically across genotypes when grown on flg22-supplemented media. *desi3a-1*, A1, A2 and B1 transgenic lines were significantly heavier than Col-0. For each genotype, seedlings grown on ABA-containing mediums were significantly lighter in comparison to seedling grown in normal conditions. The reduction in fresh weight of seedlings exposed to ABA from non-exposed seedlings was approximately 65%, 68%, 60%, 69%, 76% and 68% for Col-0, *ots1 ots2*, A1, A2, A3 and B1 respectively. The exogenous flg22 application only caused a significant difference in the A1 line. A1 seedlings grown on flg22 media were significantly heavier than seedlings grown in normal conditions. This analysis discredited the theory that flg22 promotes the increase in biomass for *desi3a-1* and A2 transgenic lines.

3.6.2.3. Root Length Assay on T3 KO Transgenic Lines

The lines with 100% resistant T3 seedlings (Table 3.4), A15-8 and A1-3 lines, were subjected to another root length assay. The same control and mutant lines were used again for comparative measures and the exact same protocol as the T2 root length assay was followed. Representative samples for each genotype on the normal and treated mediums at day 5 post-transfer are displayed in Figure 3.15-A. Quantification of the root lengths were calculated using averages of all seedlings tested per genotype, comprising of at least 10 induvial seedlings. Due to time constraints, this experiment was only repeated once. The average root lengths for each genotype grown on MS, ABA and flg22 MS media are presented in Figure 3.15-B. Statistical significance was calculated using the Kruskal-Wallis test followed by a Bonferroni post-test. To quantify the amount of root growth inhibition the stress elicitors exerted, the difference between the average root length of seedlings grown on MS media and seedlings grown on ABA or flg22 supplemented media of the same genotype was calculated. This assay presented in Figure 3.15-C discerns whether the difference in root elongation was a consequence of overall slower growth of the genotype or the presence of stress inducers restricting root growth of the seedlings. Statistical significance was calculated using the Mann-Whitney U-test which compared seedlings grown in normal and stress-induced conditions.



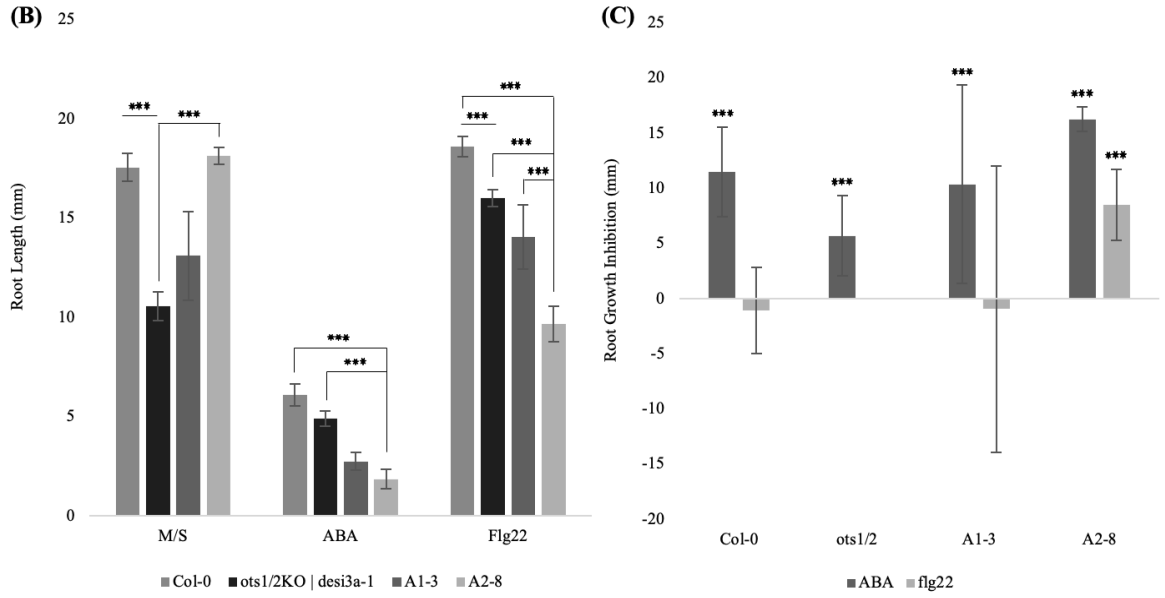


Figure 3.15 – Results from the T3 Root Length Assay. A) A representative snapshot of root elongation of each genotype grown on MS plates and ABA (25µM) and flg22 (250nm) supplemented MS plates for 6 days. Scale bar equals 1cm. B) Quantification of average root length for each genotype grown on normal and treatment containing media. C) Quantification of average root growth inhibition for each genotype comparing the difference between root lengths of seedlings grown in normal and stress-induced conditions (ABA and flg22). Data presented are means ± SE from at least 10 individual seedlings for each genotype. Significance was assessed using the Mann-Whitney U test for 2 samples and Kruskal-Wallis test followed by Bonferroni post tests for comparing 3 samples or more. Significance values: * = $P < 0.05$; ** = $P < 0.01$; *** = $P < 0.005$.

When comparing between genotypes grown on each medium, the Kruskal-Wallis test determined a statistically significant difference in mean root length between genotypes grown on all 3 mediums. Post hoc tests revealed that in normal conditions, the root length of *ots1 and ots2* was significantly shorter than Col-0 and A2-8. In ABA-induced conditions, the A2-8 mutant line was significantly shorter than Col-0 and *ots1 and ots2*. Similar results were found when comparing average root lengths between genotypes of seedlings grown on flg22-supplemented MS media. The A2-8 line was significantly shorter than all the other genotypes tested, whereas Col-0 was significantly longer than *desi3a-1* and A2-8 line. The Mann-Whitney U test for the root elongation inhibition assay found that for all genotypes, the presence of ABA significantly inhibited root growth. The most affected genotype by ABA was the A2-8 transgenic line, with an average of 16.26mm inhibition of root length compared to seedlings grown on MS media. The root length of Col-0 plants was inhibited by 11.46mm whilst *ots1 and ots2* seedlings were only inhibited by 5.66mm, implying AT60-AT80KO lines may be more sensitive to ABA. With large SE margins for the A1-3 line, the data and statistics could not be considered accurate. Exogenous application of flg22 did not affect Col-0 or *ots1 and ots2* significantly, however, the treatment did significantly inhibit root elongation in A2-8 lines. This indicates the AT60-AT80KO mutant may be more sensitive to the presence of flg22 affecting root growth mechanisms within the transgenic plant. Overall, this experiment showed how the presence of both stress elicitors inhibited root elongation in AT60-AT80KO mutants, suggesting the KO lines could be significantly more sensitive to ABA and flg22 than WT.

3.7. Discussion

This extensive results chapter first looks into defining the genetic and proteomic characteristics of the AT60 and AT80 DeSi SUMO proteases. Bioinformatic, protein modelling and gene expression analyses of the two proteases were performed using online resources and software programs. Experimental genetic analyses and preliminary subcellular localisation studies in *N. benthamiana* of AT60 and AT80 were conducted in the laboratory to evaluate findings from online resources and provide further information on the two DeSi SUMO proteases. The latter half of this chapter describes the generation of the AT60-AT80KO mutant plants with the CRISPR/Cas9 system, and the subsequent genotyping tests performed on the mutants to validate their mutagenic attributes. Lastly, the transgenic plants were subjected to phenotypic assays in normal and stress-induced conditions to investigate the function of the AT60 and AT80 DeSi proteases in the development and defence system of Arabidopsis plants.

3.7.1. Bioinformatics and proteomic analyses reveal AT60 and AT80 proteases are strikingly similar to one another

Before working on the DeSi SUMO proteases, online resources and software programs were used to research and define the genetic and proteomic characteristics of AT60 and AT80. The amino acid sequence alignments of the two proteins revealed they share 93% identical amino acids, which was a strong indication that the proteases would have very similar protein structures. The proteomic analyses predicting the 3D molecular structure of AT60 and AT80 confirmed this notion. Furthermore, the gene expression profiles determined by the ePlant browser also displayed similarities in the spatial pattern and expression levels of the respective genes. These findings confirmed the two DeSi proteases share very similar characteristics and are, therefore, most likely a result of a gene duplication event. This indicates there could be a strong probability that the proteins are genetically redundant. This can be seen to an extent, in other Arabidopsis SUMO proteases including OTS1/2 and SPF1/2.

OTS1 and OTS2 proteases form a key subgroup within the ULP1 SUMO protease group. Phylogenetic analyses found the two SUMO proteases to be highly homologous to one another sharing 56% amino acid identity. When comparing only within the protease domain, the amino acid identity increased to 73% (Conti *et al.*, 2008). The two proteins have been shown to localise to the nucleus and function partially redundantly in regulating salt stress responses (Conti *et al.*, 2008), flowering transition, light-induced signalling (Sadanandom *et al.*, 2015) and have been implicated in SA signalling in Arabidopsis plants (Bailey *et al.*, 2016). The genetic redundancy between OTS1 and OTS2 was discovered when *ots1 ots2* exhibited increased sensitivity to salt stress in comparison to single *ots1* and *ots2* mutants (Conti *et al.*, 2008). Similarly to OTS1/2, the AT60 and AT80 DeSi SUMO protease share very high amino acid identity and in fact, were even more homologous to one another than OTS1/2. Therefore, it was necessary for the KO transgenics of AT60 and AT80 to be double mutants as the likelihood of the two proteases being functionally redundant was very high. On the other hand, the SPF1/2 SUMO proteases which form a key subgroup within the ULP2 SUMO protease group, only share 30.5% identity; significantly lower than the amino acid identity between AT60 and AT80. Although the double KO of SPF1/2 resulted in a diverse number of mutant phenotypes including late flowering, altered leaf morphology and impaired seed production, reports from various studies support the existence of unequal redundancy. Findings support the notion that SPF1 was the more important SUMO protease as SPF1 is

expressed more and the *spf1* mutant showed more notable phenotypes in leaf morphology, flowering time, and pigment accumulation as well as greater SUMO conjugate levels (Castro *et al.*, 2018). Therefore, the possibility of unequal redundancy between the AT60 and AT80 protease should also be considered. To substantiate this theory, the generation and analyses of AT60 and AT80 single KO mutants should have been performed and tested, if more time was available.

3.7.2. Genetic analyses and subcellular localisation studies of AT60 and AT80 provides some indication of their localisation and function in Arabidopsis

Research into the two DeSi SUMO proteases began with genetic analyses. First, the presence of the mRNA transcript of the AT60 and AT80 gene was tested to confirm the expression of both proteases in Col-0 plants. Subsequently, a real-time PCR was conducted to determine the expression levels of AT60 and AT80 across various plant organs and developmental stages of Col-0 plants, as well as their expression in response to ABA and flg22. This experiment would provide some indication on the localisation and function of both DeSi proteases, identifying whether there are discernible differences in their gene expression profiles, as well as to corroborate the findings from the online ePlant tool.

Overall, the results from the qPCR showed the gene expression level and spatial expression patterns of AT60 and AT80 were very similar. However, AT80 expression levels were greater in the stem and roots of the plant, whilst AT60 expression was higher in the cauline leaves. The comparative analysis on the ePlant browser showed higher expression levels of AT60 across all plant organs and developmental stages of Arabidopsis, except in mature pollen where AT80 expression was considerably greater. The only qPCR result that aligned with the ePlant browser findings was the higher expression levels of AT60 in the cauline leaves of Col-0 plants. The qPCR performed on 13-day old seedlings subjected to stress elicitors resulted in notable differences in the AT60 and AT80 expression levels. In response to ABA, the transcript level of AT80 increased five times more than seedlings in normal conditions, whilst AT60 expression decreased by 48%. In the presence of flg22, both AT60 and AT80 transcript levels considerably reduced. This experiment provided initial insights into the possible role the DeSi proteases may have in modulating ABA and immune signalling in Arabidopsis. Subsequent phenotyping assays further analysed this association. The qPCR results provided more information on AT60 and AT80 expression in response to stress than the findings from the ePlant browser which found no changes in transcript levels of either proteins when subjected to biotic or abiotic stresses. However, having only conducted three technical repeats, limitations to the findings from the qPCR experiments should be considered and not regarded as conclusive.

The most probable explanation for the differences observed between published and experimental expression of target genes is the significant variation in environmental conditions. However, technological differences could also be considered. The data populating the ePlant browser was obtained from studies that conducted microarray analyses using the ATH1 Affymetrix Arabidopsis GeneChip on Arabidopsis plants and seedlings (Nakabayashi *et al.*, 2005; Schmid *et al.*, 2005). Microarray analyses are a powerful tool for studying an entire genome of an organism in a single experiment, particularly the ATH1 Affymetrix GeneChip, which contains over 22,000 redesigned probe sets spanning most of the identified cDNA and open reading frames of Arabidopsis (Hennig *et al.*, 2003). However, as ATH1 arrays only comprise of probes for known genes

available as of 2001, any new discoveries or information of recent genome versions are not included in the probe design (Olbricht *et al.*, 2009). Furthermore, with the mass number of genes analysed at once, data result interpretation can be difficult, individual hybridisations can be noisy leading to reproducibility issues and single data points may be unreliable. This can especially be seen with genes that are low in abundance (Casneuf *et al.,* 2007). Therefore, although the results from the ePlant browser can provide some insight into the gene expression profile of AT60 and AT80, the differences observed between results from the qPCR experiment and ePlant tool may be due to the issues associated with the ATH1 Affymetrix Arabidopsis GeneChip in addition to the variation in environmental conditions.

The expression patterns of the Arabidopsis proteases, SPF1/2, were established using two different approaches; a semiquantitative RT-PCR and β-glucuronidase (GUS) staining for SPF1/2 promoter analysis. A RT-PCR of SPF1/2 in different tissues of flowering Arabidopsis plants determined the presence of the two proteases in most plant structures, with greater expression levels in the developing reproductive organs such as flower buds and siliques, as well as in the cauline leaves. GUS-fused SPF1 and SPF2 promoters were introduced into Col-0 plants and the GUS activity was identified primarily in the floral organs and developing embryos of the transgenic plants (Liu *et al.*, 2017). These two approaches were efficient in determining the expression pattern of SPF1/2 and therefore may have been a more accurate method for discerning the expression profile of AT60 and AT80.

Lastly, a transient expression assay of AT60 and AT80 in *N. benthamiana* plants was conducted to provide an indication on the subcellular localisation of the two DeSi proteases. The immunoblot analyses proved the recombinant AT60 and AT80 proteins were successfully fused to the fluorescent and epitope tags. This confirmed the recombinant bacterial strain overexpressing AT60 and AT80 efficiently expressed the respective recombinant proteins and would therefore be critical for subsequent experiments. In general, studied SUMO proteases in Arabidopsis have been observed to localise primarily in the nucleus as seen with the SPF1/2 and OTS1/2 proteases (Conti *et al.*, 2009). The only known SUMO protease to localise outside the nucleus in the cytosol is ELS1 (Hermkes *et al.*, 2011). Very recently, a transient expression assay found the DeSi SUMO protease, Desi3a, to localise to the plasma membrane (Orosa *et al.*, 2018). Similarly, in this confocal microscopy study, both AT60 and AT80 proteins were found to localise along the cell periphery, most likely the plasma membrane and faintly in the nucleus. However, the expression of both recombinant proteins was very strong being driven by the CaMV35S promoter, and as the experiment lacked the essential cellular markers, the compartmentalisation of the proteases could not be concluded. Instead of driving the expression of the genes with the CaMV35S constitutive promoter, cloning the AT60 and AT80 promoter with the respective genes would have resulted in more accurate findings. Although this was attempted throughout the research duration, the overlapping PCR protocol was unsuccessful and therefore could not be achieved within the limited time period. Confocal microscopy studies in the AT60 and AT80 overexpressing transgenics lines were subsequently performed and the findings are detailed in section 4. This experiment provides more conclusive results on the subcellular localisation and spatial expression pattern of the AT60 and AT80 DeSi proteases across the Arabidopsis seedling.

3.7.3. The efficiency of the CRISPR/Cas9 system to generate heritable homozygous AT60-AT80KO mutants was assessed

As gene KO studies are fundamental in elucidating the function of a specific gene of interest, the state-of-the-art targeted mutagenesis tool, the CRISPR/Cas9 system, was utilised to generate AT60-AT80KO mutant Arabidopsis plants. The CRISPR/Cas9 protocol performed in this study was derived from the strategy established by Xing *et al.* (2014) and Wang *et al.* (2015). A CRISPR/Cas9 binary vector set containing multiple sgRNA expression cassettes was used to generate the KO transgenics. An expression cassette with two sgRNAs targeting AT60 was transformed into homozygous AT80 T-DNA KO mutants with confirmed T-DNA insertions disrupting the AT80 gene. An expression cassette with sgRNAs targeting the AT60 and AT80 gene was also introduced into Col-0 plants. These two approaches were used to generate AT60-AT80 double KO mutant plants. The recombinant *Agrobacterium* carrying each of the sgRNA expression cassettes were transformed into the respective Arabidopsis background plant using the floral dipping method. The T1 seedlings were screened on hygromycin MS selection plates and the resistant seedlings were subsequently grown for genotyping analyses and propagation.

To screen for homozygous mutants in the T1 generation and evaluate mutation efficiencies, purified PCR fragments flanking the target sites were sequenced. PCR fragments were isolated from genomic DNA extracted from rosette leaves of T1 AT60KO in AT80KO (A1-A3) and AT60-AT80KO in Col-0 (B1) transgenic lines. The sequencing results were aligned with the WT sequence of the corresponding target region. This analysis identified mutations had occurred in the expected targeted DNA region located near the DSB site in all tested T1 transgenic lines. Analysing the chromatogram peaks of the DNA sequences for each T1 mutant line, confirmed all identified indel mutations generated by CRISPR-Cas9 to be homozygous. However, as genomic sequencing depends entirely upon the quality of the purified PCR fragment and the analysis was only repeated once, the results are prone to misalignment, miscalls and human error and therefore, inconclusive. Due to limited evidence, the sequenced T1 transgenic lines could only be presumed to be homozygous CRISPR-Cas9 KO mutants and the possibility of unsuccessful transformation was taken into consideration in later experimental steps. As time was limited, the CRISPR-Cas9 protocol and sequencing analysis could not be repeated, and the T1 transgenic lines were propagated to the T2 and T3 generation for phenotyping analyses and to further establish the zygosity of the potential mutants. However, as the uncertainty in the potential T1 transformant lines undermines the legitimacy of the T2 and T3 KO mutants, results from subsequent phenotyping analyses should be interpreted carefully and not considered conclusive.

The study by Wang *et al.* (2015) proved that by using egg cell-specific promoters to drive Cas9 expression, non-mosaic T1 Arabidopsis mutants for multiple target genes could be generated. In this paper, phenotypic segregation was used to screen for homozygous mutants. As the three target genes *ETC2*, *TRY*, and *CPC* had highly recognisable phenotypes (clustered leaf trichomes) when knocked-out, this study could screen for T1 homozygous mutants without genotypic analyses. However, in most cases including this experiment, there was no convenient or visible phenotype for the genes-of-interest. Direct sequencing of the PCR fragment spanning the target region works efficiently and accurately as a screening method, especially for homozygous or bi-allelic mutations with a single base pair insertion, deletion or substitution. However, this can be a very expensive method particularly when screening a larger sample size. Alternatively, a primary screening step could be conducted with a restriction enzyme digestion analysis or a T7E1/Surveyor assay, which detects endogenous target cleavages (Cong *et al.*, 2013). Fortunately, in this experiment, the sample size of the T1

transgenic lines was relatively small and therefore, sequencing was the most time-efficient method for screening homozygous mutants and evaluating mutation efficiencies.

The results obtained from the sequencing analyses aligned with findings from published papers. CRISPR/Cas9 activity produces mutations with only a few nucleotide additions or deletions within the sgRNA target region as a consequence of the error-prone NHEJ pathway (Cho *et al.*, 2017). As the Cas9 endonuclease cleaves one single strand of the double-stranded DNA approximately three-nucleotides upstream from the PAM site, all indels would be expected to be situated in close proximity to the cleavage site (Jiang *et al*., 2014). This has been observed in studies that have successfully utilised the CRISPR/Cas9 system to KO their gene(s) of interest including the generation of cold-responsive *C-repeat/DRE-Binding Factor* (*CBF*) mutants (Cho *et al.*, 2017) and successfully converting a mutant green fluorescent protein (GFP) gene to a functional GFP gene using Cas9/sgRNA-directed mutagenesis (Jiang *et al.*, 2014). In this experiment, only two T1 transgenic lines, A1 and B1, had indels located in close proximity to the expected cleavage site. However, as these mutations were still situated several nucleotides away from the expected PAM site, they could not be conclusively regarded as genuine AT60-AT80KO mutants. Furthermore, the mutations identified in the T1 transgenic lines further upstream from the target site were more likely attributable to the low-quality start of the electropherogram instead of the CRISPR-Cas9 system. If more time were available, the screening analyses would have been repeated several times and more KO transgenic lines would have been generated to be screened. As many of the identified mutations in all T1 transgenic lines were not within a few bases of the PAM site as expected for this method, it could only be inferred and not confirmed that the mutations observed in all T1 transgenic lines were induced by the CRISPR-Cas9 system. Therefore, the possibility of unsuccessful transformation and therefore the unsuccessful creation of AT60-AT80KO mutants at the T1 generation of should be considered in later experimental stages.

To ensure the mutations produced by CRISPR/Cas9 were germline inherited, the binary vector system developed by Wang *et al*. (2015) and Xing *et al.* (2014) was selected. The Cas9 endonuclease was driven by the promoter from the egg cell-specific *EC1.2* of Arabidopsis to induce mutations in the egg cells thereby enhancing heritability. Studies using this binary vector system were successful in generating heritable genetic mutations in the T2 and T3 progeny. For instance, a study by Wolter *et al*. (2018) used the same binary vector system to produce an extensive data set obtaining over 1000 heritable gene targeting events within the T2 generation. Therefore, with this information and the inaccuracy in validating homozygosity of indel mutations in the T1 transgenic lines, the mutants were propagated to the T2 and T3 generation prior to phenotyping analyses. Screening of T2 and T3 seedlings on hygromycin MS selection plates was performed and resistant seedlings were either subjected to phenotyping analyses or propagated to the next generation. The T3 lines with 100% resistant seedlings could be considered homozygous double KO mutant lines. This is because CRISPR/Cas9 activity should continue within T1 egg cells, T2 one-cell stage embryos and T2 early embryos. Therefore, T1 plants that may be heterozygotes or mosaics would be able to give rise to homozygous or bi-allelic mutant T2 or T3 plants (Wang *et al.,* 2015). The T2 and T3 AT60-AT80KO mutant lines would have been sequenced again to validate the mutations were inherited in the progeny lines if more time was available. However, results from the sequencing and segregation analyses determined which T2 and T3 AT60-AT80KO lines could be considered genuine homozygous AT60-AT80KO mutants and were therefore subsequently subjected to phenotyping analyses.

3.7.4. Phenotyping analyses of the AT60-AT80KO mutants illustrate a role for AT60 and AT80 in the stress signalling pathways of Arabidopsis plants

The T2 and T3 AT60-AT80KO mutant lines were phenotyped with a root length and fresh weight assay to assess whether the mutants display pleiotropic phenotypes under normal conditions and in response to biotic and abiotic stresses. However, due to time constraints, the fresh weight assay was only conducted on the T2 mutant lines. The key findings from all three phenotypic analyses were considered and analysed together in this section. However, as the zygosity of the AT60-AT80KO mutant lines analysed were still uncertain at the T2 generation, the results from the T3 root length assay could be considered more of an accurate representation of the phenotypes observed in AT60-AT80 KO mutant plants. Across all three phenotypic analyses, the key findings were as follows.

In the root length assay on the T2 transgenic lines, under normal conditions, there was no statistically significant difference in root growth across all tested genotypes. The fresh weight assay was in alignment with these results as all genotypes weighed similar amounts to one another. However, in the T3 root length assay, the *ots1 ots2* seedlings had significantly shorter root lengths than WT and one of the AT60-AT80 KO mutant lines (A2-8). In the published literature, no significant difference in root elongation rates between *ots1 ots2* and WT seedlings have been reported (Conti *et al.*, 2008). Therefore, it can be concluded that in normal conditions, no significant differences were observed in the root growth or biomass production between the SUMO protease KO mutants and Col-0.

For the stress response root length and fresh weight assays, the phenotypes were compared between each of the genotypes analysed including Col-0 and other SUMO protease KO mutants, as well as between seedlings grown in normal and stress-induced conditions within each genotype. This was conducted to determine whether the phenotype observed was driven solely by the presence of the stress elicitor, ABA or flg22, and was synonymous with the WT phenotype, or if the observed phenotype was unique to the genotype of the AT60-AT80KO mutants.

In the T2 root growth assay, the presence of ABA significantly inhibited the root growth of one AT60-AT80KO mutant line (B1) and *ots1 ots2*, whilst in the other three mutant lines, root elongation decreased by only a small amount. As B1 also had significantly shorter root lengths in comparison to Col-0, this AT60-AT80KO mutant line was considerably more susceptible to ABA than WT. In the T3 root length assay, significant root growth inhibition occurred in all genotypes, with the AT60-AT80 KO mutant (A2-8) exhibiting the greatest inhibition. As the A2-8 line also had significantly shorter root lengths than Col-0 and *ots1 ots2*, this finding further supports the notion that ABA impedes root development in AT60-AT80KO mutants considerably more than WT, and therefore AT60-AT80KO mutants could be hypersensitive to ABA. The T2 fresh weight assay found that the biomass of all genotypes significantly decreased in response to ABA, with the greatest reduction denoted in the AT60-AT80KO line, A3. However, as no difference was detected in the biomass across all genotypes, this phenotype was not exclusive to A3. As these results were consistent with the T3 root length assay, ABA most likely triggers the switch from growth to defence mode in the SUMO KO mutants suppressing global growth of all plant tissues including root development. Overall, these experiments confirmed that in the presence of ABA, the AT60-AT80KO mutant lines displayed significant root growth inhibition relative to WT

implying these mutant lines are more sensitive to abiotic stresses. The biomass of AT60-AT80KO mutants also reduced with ABA application, however, not to a significant degree in comparison to WT.

In the T2 root length assay, the application of flg22 enhanced the root growth in Col-0 and the three AT60-AT80KO mutant lines. However, as there was no statistical difference between root lengths of WT, *desi3a-1* and the T2 KO mutant lines, the increased root growth observed in response to flg22 was not distinctive to the AT60-AT80KO mutant lines. On the contrary, in the T3 root length assay, significant root growth inhibition occurred in the AT60-AT80KO mutant line, A2-8. In comparison to the other genotypes, this A2-8 line also had significantly shorter roots relative to WT and *desi3a-1* seedlings when exposed to flg22. Even the *desi3a-1* lines had reduced root elongation in comparison to Col-0. As the T3 AT60-AT80KO mutant lines were more likely to be genuine homozygous AT60-AT80KO mutants, the three T2 AT60-AT80 KO mutant lines displaying the same phenotype as the WT could be considered non-mutants. Instead, the findings from the T3 assay suggested the AT60-AT80KO mutants were considerably more sensitive to flg22 than WT, resulting in the inhibition of root elongation. As the fresh weight assay was only conducted on the T2 transgenic lines which were considered non-mutants, the results from this assay could not be taken into consideration. Overall, this experiment found that flg22 considerably impedes the root growth of AT60-AT80KO mutants and are therefore more sensitive to flg22 than WT.

The results from these phenotyping experiments suggest the AT60-AT80KO mutants are more sensitive to the presence of both ABA and flg22 relative to WT. These findings are consistent with published studies on the response SUMO protease KOs have to biotic and abiotic stresses. The Arabidopsis SUMO protease, OTS1/2 double mutant displayed extreme sensitivity to salt stress in comparison to WT or single *ots1* and *ots2* mutants. As salt stress is one of the primary abiotic stresses in Arabidopsis, the findings from this experiment are consistent with the study by Conti *et al.* (2008). On the contrary, AtSIZ1 has been shown to negatively regulate ABA signalling as *siz1* mutants display ABA hypersensitivity resulting in germination and primary root growth inhibition (Miura *et al.*, 2010). Although similar results have been shown in the *ots1 ots2* mutants in relation to salt stress, previous studies have reported that the triple-mutant *ots1/2 siz1* displayed accumulative defects, which place OTS1/2 and SIZ1 on separate pathways (Castro *et al.*, 2016). As similar results to *ots1 ots2* in response to salt stress were identified in this study, it could be postulated that the two DeSi proteases are also on different pathways to the AtSIZ1 enzyme. Therefore, concluding that AT60 and AT80 could be involved in negatively regulating ABA signalling.

Flg22 of bacterial flagellin is recognised by the leucine-rich repeat (LRR) receptor kinase, FLS2, during bacterial infections to induce pattern-triggered immunity (PTI). In normal conditions, FLS2 is complexed with an intracellular kinase, BOYTRISTIS INDUCED KINASE1 (BIK1). When FLS2 detects flg22, the kinase recruits and dimerises with the transmembrane kinase protein BRASSINOSTEROID INSENSITIVE1 ASSOCIATED KINASE1 (BAK1) to directly phosphorylate BIK1 resulting in its dissociation from FLS2 which in turn activates downstream signalling components to induce PTI (Igarashi et al., 2012). Recently, it was established that the mechanism governing the dissociation of FLS2-BIK1 depends on the SUMOylation of FLS2 triggering the release of BIK1 (Orosa et al., 2018). Disruption of FLS2 SUMOylation can abolish immune responses, resulting in susceptibility to bacterial pathogens in Arabidopsis. The most common method to investigate FLS2-dependent antibacterial immunity has been to conduct root length assays testing for flg22-

mediated root growth inhibition. Recent studies established that roots are able to perceive flg22 and trigger the PTI defense mechanism (Chuberre et al., 2018). Flg22 treatment in Arabidopsis increases resistance to microbial invasion in roots during PTI response through promotion of reactive oxygen species (ROS) accumulation, callose deposition and development of antimicrobial compounds, strong tissue-specific upregulation of defence-related responses and accumulation of the defence hormone SA (Millet et al., 2010). As a result of the activation of energy-costly defence mechanisms, exposure to PAMPs such as flg22 have a negative effect on plant growth including root growth inhibition.

The paper by Orosa *et al.* (2018) was able to establish a regulatory role for the Arabidopsis DeSi protease, Desi3A, in immune signalling through conducting flg22-induced root growth assays. This study found that upon flg22 perception, the Desi3A protein rapidly degrades leading to hyper-SUMOylation of FLS2 which activates immune signalling in Arabidopsis plants. The involvement of Desi3a in flagellin sensing was supported by flg22-induced root growth assays performed on Col-0 and *fls2* and *desi3a-1* mutants. The presence of flg22 induced significant root growth inhibition in Col-0 whilst *fls2* exhibited less sensitivity to flg22, with significantly greater root elongation than Col-0. *desi3a-1* root elongation was found to be more sensitive to flg22 than Col-0 due to the complete lack of regulation of FLS2 SUMOylation. This was consistent with the results of AT60-AT80KO mutants in the root length assay. Therefore, these findings suggest that AT60 and AT80 also play a role in flagellin sensing and are likely to be negatively regulating plant immune responses and inducing root growth inhibition in the early stages of a pathogen infection like Desi3A. The results from this experiment suggest a role for AT60 and AT80 in the abiotic stress and immune signalling pathways of Arabidopsis. These findings are further explored in the next chapter where the same phenotyping analyses were performed on AT60 and AT80 overexpressing plants.

3.7.5. Future work in relation to this study

There are various factors and experiments that would have been performed to contribute towards characterising and defining the role of the AT60 and AT80 DeSi proteases. Firstly, the generation of single AT60 and AT80 KO mutants and subsequent phenotyping analyses would have determined whether the DeSi proteases are functionally redundant or partially redundant. Furthermore, this would substantiate the qPCR results on whether AT60 functions primarily in the cauline leaves, which was where AT60 was highly expressed, and if AT80 plays a significant role in the stem and roots; the plant organs where higher AT80 transcript levels were detected. A more accurate method in determining homozygosity in the AT60-AT80KO mutants would have been to sequence all T2 transgenic lines again to identify homozygous indel mutations. Those carrying homozygous mutations would be propagated to T3 resulting in the generation of a genuine homozygous AT60-AT80KO mutant. Furthermore, with more time it could have been possible to segregate out the Cas9 enzyme in the T2 generation or through propagating heterozygous CRISPR-Cas9 mutants. This could have been achieved by selecting and growing the non-resistant seedlings when screening for transformants on selection plates, then subsequently confirming the presence of the mutations at the locus of interest. The most promising transgenic lines would be propagated to the next generation, resulting in the creation of stable Cas9-free homozygous AT60-AT80KO mutant lines (Pauwels *et al.*, 2018). More phenotyping experiments on the AT60-AT80KO mutant lines would have been performed including germination rate assays and flowering time experiments, to further elucidate the role of AT60 and AT80 in Arabidopsis development. Lastly, genetic

complementation tests in the AT60-AT80KO mutants should have been performed to confirm the phenotypes observed in the mutants were caused by the AT60 and AT80 DeSi protease.

3.7.6. Final concluding remarks

This chapter first details how the two AT60 and AT80 SUMO proteases share very similar characteristics to one another highlighting the possibility the two proteases resulted from a gene duplication event. With this information, it was postulated that AT60 and AT80 function redundantly. The genetic analyses performed established similarities and differences in the gene expression profiles of both DeSi proteases. For instance, AT60 was greatly expressed in cauline leaves, whilst AT80 was more abundant in the stem and roots of Arabidopsis plants. The response of AT60 and AT80 to flg22 was similar as both significantly reduced in transcript levels. In response to ABA, their expression levels diverged; AT60 marginally decreased, whilst AT80 considerably increased. The subcellular localisation study in *N. benthamiana* provided some indication that both proteases primarily localise along the cell periphery, most likely the plasma membrane and faintly in the nucleus. This chapter then detailed how the homozygous AT60-AT80 KO mutant lines were generated using the CRISPR/Cas9 binary vector system and the subsequent genotyping experiments that evaluated the mutagenic attributes and zygosity of the mutant lines. The root length and fresh weight assay in response to normal and stress-induced conditions found that AT60 and AT80 may be implicated in negatively regulating ABA and flg22 signalling pathways in Arabidopsis. This suggests that both DeSi SUMO proteases could be involved in the development and defence system of Arabidopsis plants. These findings were further substantiated in the next section.

4. Generation and Analysis of AT60 and AT80 Overexpressing Transgenic Plants

4.1. Introduction

Traditionally in genetic analyses, the investigation and elucidation of biological pathways usually begin with the identification of mutations giving rise to a phenotype of interest. However, overexpression of a WT gene which leads to the abundant expression of the target protein can also result in mutant phenotypes. Therefore, overexpressing genes of interest can act as an alternative yet effective tool for studying the characteristics of pathway components which may not be identified using conventional loss-of-function analysis. For instance, essential genes and redundant members of gene families are the types of genes which confer only subtle phenotypes when knocked-out and require highly specific alterations in the protein (LeClere and Bartel, 2001).

As previously mentioned, elaborate genetic analyses of mutants including gain-of-function transgenics, of the SUMO enzymatic machinery have significantly contributed to the study of the SUMO system. The paper by Lois *et al*. (2003), exemplified a role for SUMO in modulating the ABA signal transduction pathway. The study demonstrated how transgenic plants overexpressing AtSUMO1/2 hindered ABA-mediated growth inhibition and enhanced the expression of ABA- and stress-responsive genes. Another paper which investigated the AtSUMO1/2 paralog AtSUMO3 found that overexpression of AtSUMO3 resulted in early flowering and plant defence activation. This overexpression study was able to determine that AtSUMO3 plays a role in promoting plant defences downstream of SA, whilst AtSUMO1 and 2 jointly hinder SA accumulation in non-infected plants (van den Burg *et al*., 2010). When investigating SUMO proteases, overexpressing OTS1 showed greater resistance to high salinity conditions, indicating the OTS1 protease activity promotes salt tolerance by reducing SUMOylation levels (Conti *et al.*, 2008). Therefore, overexpression studies have been critical in elucidating the mechanism of the SUMO system.

To further analyse the AT60 and AT80 DeSi proteases and identify characteristics which may not have been identified using the loss-of-function analysis, transgenic plants overexpressing either of the DeSi proteases were generated. Single overexpressing transgenics were generated to determine whether the two highly homologous proteases function redundantly or divergently. If more time were available, double overexpressing AT60 and AT80 transgenics would have been produced as a comparative measure to the single gain-of-function mutants. This chapter explores how the single overexpressing transgenic plants were generated using the pEG vector, and subsequently presents results from the phenotyping analyses performed on the overexpressing transgenics to determine the impact AT60 and AT80 DeSi proteases may have on Arabidopsis plant development.

4.2. Constructing Overexpressing Transgenic Plants

4.2.1. Introduction to the Construct and Transformation

The most commonly used strong constitutive promoter to drive high levels of gene expression and generate overexpressing transgenic plants is the CaMV 35S promoter. pEG vectors utilise the enhanced CaMV 35S promoter for strong constitutive expression of proteins fused to a variety of oligopeptide epitope tags and fluorescent proteins, such as GFP or YFP. These gateway cloning-compatible destination vectors are not only

useful for protein overexpression studies but can also be used for immunoblotting, affinity purification and subcellular localisation experiments (Earley *et al.*, 2006).

For the generation of single AT60 and AT80 overexpressing Arabidopsis plants, the pEG101 vector with YFP and hemagglutinin (HA) C-terminal tags was used. Dr. Orosa cloned the cDNA gene fragments of AT60 and AT80 into pEG101 and transformed the recombinant vector into competent *Agrobacterium* cells. Glycerol stocks of AT60 and AT80 overexpressing *Agrobacterium* cells were kindly donated by Dr. Orosa for the purpose of this experiment. For the delivery of the overexpressing recombinant vectors into Col-0 plants, the floral-dipping method was used. The dipped plants were subsequently grown in normal conditions and all seeds were harvested for mutant screening.

4.2.2. Screening T1 Transgenic Plants and 3:1 Segregation Analysis of T2 Transgenic Plants

The transgenic seeds harvested from the dipped T0 plants were resistant to the antibiotic BASTA if successfully transformed. Therefore, positive selection was achieved by growing the seeds on soil watered with 1:1000 BASTA to water mix. Seedlings were screened for transformants and those which grew normally were allowed to mature in normal conditions. Figure 4.1-A displays the AT60 and AT80 BASTA-resistant seedlings which were selected to grow to maturity. Seeds were then harvested from the T1 transgenic plants and assayed by growth on BASTA (30µg mL^{-1}) MS selection plates. The results from this T2 segregation analysis are displayed in Figure 4.1-B. T2 resistant seedlings from T1 lines producing seedlings at a 3:1 ratio of resistant to non-resistant seedlings, were grown in normal conditions and subjected to various genotyping analyses.

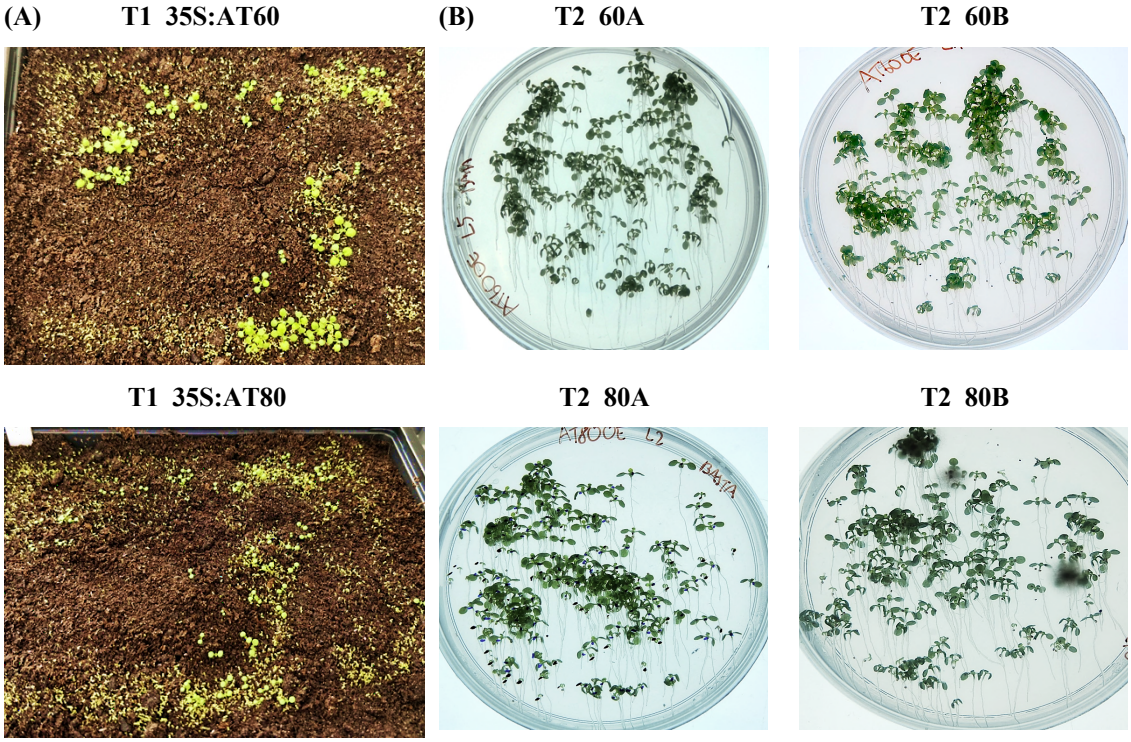


Figure 4.1 – Screening T1 overexpressing transgenic seedlings and 3:1 segregation analysis of T2 seedlings. A) Seeds harvested from *Agrobacterium*-transformed T0 plants were sown on soil watered with 1:1000 BASTA to water mix. Resistant seedlings were transferred to normal soil and grown for seed harvesting. B) T2 transgenic seeds harvested from T1 transgenic lines were grown on BASTA (30µg mL^{-1}) MS selection

plates and the lines segregating at a 3:1 ratio were grown to maturity. Lines displayed are representative samples of all T2 lines analysed.

4.3. Genotyping the T2 Overexpressing Transgenic Lines

Various genotyping experiments were conducted to confirm the newly generated overexpressing transgenic plants were truly gain-of-function mutants. Firstly, the RT-PCR checked whether the T2 overexpressing plants were carrying the recombinant pEG101 vector containing the YFP- and HA-tagged recombinant protein. qPCR was performed on the cDNA extracted from both T2 and T3 overexpressing transgenic lines to assess whether an increased expression of the overexpressed DeSi protease could be observed. Lastly, to confirm the presence and enhanced expression of the transgene in the T2 overexpressing transgenic lines, total protein extracted from the overexpressing lines were analysed by SDS-page and immunoblotting. The results from these experiments confirmed which overexpressing transgenic lines were carrying and overexpressing the transgene, therefore ensuring the correct T2 gain-of-function mutants were propagated to the T3 generation for phenotyping.

4.3.1. RT-PCR of T2 overexpressing transgenic lines

The RT-PCR was conducted on the T2 overexpressing lines that segregated at 3:1 on BASTA (30µg mL^{-1}) MS selection plates. cDNA was synthesised from the genomic DNA extracted from three-week old T2 overexpressing plants. To identify the recombinant vector in the gain-of-function mutants, specific primers to isolate the YFP coding sequence were used alongside actin primers as a positive control. cDNA extracted from Col-0 was also tested as a negative control. PCR reactions isolating the actin and YFP protein were conducted on the cDNA of all T2 AT60 and AT80 overexpressing lines as well as Col-0. PCR products were run on gel electrophoresis and the results of representative samples are displayed in Figure 4.2. The expected band size for the fragment isolating actin at approximately 350bp can be seen across both Col-0 and the overexpressing lines. Multiple bands present in the actin controls could be a result of off-target gene amplification. The YFP isolating fragment at approximately 200bp can also be observed in the AT60 and AT80 overexpressing transgenic lines. As expected, the YFP band cannot be observed in Col-0. This RT-PCR confirmed which T2 overexpressing lines were successfully transformed with the recombinant vector and were therefore selected for further genotyping analyses.

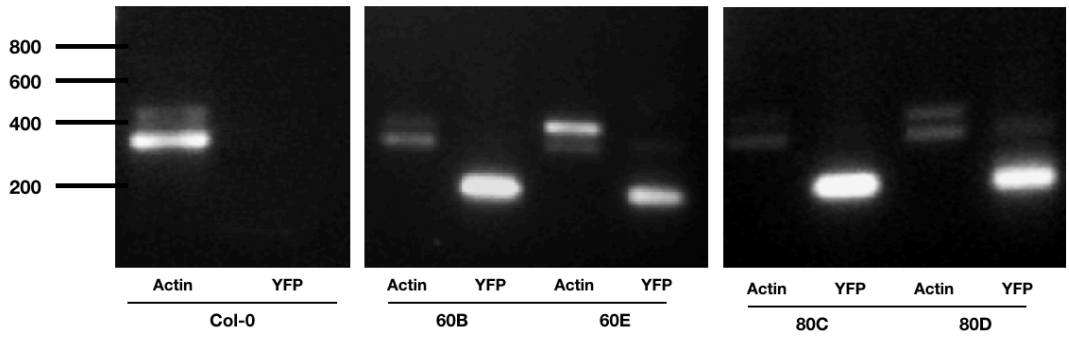


Figure 4.2 – RT-PCR results confirming T2 AT60 and AT80 overexpressing lines carry the recombinant pEG101 vector. The cDNA of three-week old Col-0 Arabidopsis plants and T2 AT60 and AT80

overexpressing transgenic plants were synthesised. For the RT-PCR, primers isolating Actin (Actin-Fwd + Actin-Rev) and YFP (YFP-Fwd + YFP-Rev) were used on the cDNA of Col-0 and T2 transgenic lines. Only the representative lines of the gain-of-function mutants were displayed. Multiple bands present in the actin controls could be a result of off-target gene amplification.

4.3.2. Real-time PCR of T2 and T3 AT60 and AT80 overexpressing transgenic lines

The expression levels of AT60 and AT80 were measured by real-time PCR in the overexpressing lines. This was conducted on both T2 and T3 gain-of-function mutants. Expression levels of both DeSi proteases were measured in the single overexpressing lines to determine whether the transcript levels of the respective transgene had significantly increased, as well as to assess whether this affected the expression level of the other DeSi protease. For the T2 generation, the lines which showed a significant increase in the overexpressed transgene were selected for T3 propagation. The T3 gain-of-function mutant lines were analysed to confirm the transgenic plants were overexpressing the respective DeSi protease and the lines with the most significant increase were selected for phenotyping analyses. For the qPCR experiment, the actin gene was used as the housekeeping gene for calibration. Technical repeats were conducted in triplicates, however, as the experiment was only repeated once due to time constraints, no statistical tests were conducted as a result of the limited sample size. To prevent misreads of actin from off-target genes as seen in the RT-PCR results, higher annealing temperatures than the RT-PCR were used in the real-time PCR. cDNA was synthesised from RNA extracted from three-week old Col-0 and T2 overexpressing lines, and 7-day old Col-0 seedlings and T3 overexpressing plants.

qPCR results of the T2 transgenic seedlings presented in Figure 4.3-A, highlights which transgenic lines have heightened expression levels of the overexpressed DeSi protein relative to WT. For the AT60 overexpressing plants, all lines except for 60C have significantly increased AT60 expression levels with over a tenfold increase observed in 60A, 60B and 60E transgenic lines. Therefore, these lines were regarded as genuine AT60 gain-of-function mutants which were for T3 propagation. Interestingly in these lines, the expression level of AT80 decreased relative to WT in the 60B and 60E lines, whilst increasing in 60A. For the AT80 overexpressing lines, all lines had heightened expression of AT80 in comparison to Col-0. Lines 80A, 80C and 80D had double the expression level of AT80 relative to WT and were therefore considered genuine AT80 gain-of-function mutants. Interestingly, heightened AT80 transcript levels also resulted in an increase in AT60 levels relative to WT. The three T2 AT60 and AT80 lines validated as overexpressing plants were considered for T3 propagation.

The real-time PCR results on the T3 AT60 and AT80 overexpressing lines are displayed in Figure 4.3-B. These selected transgenic lines were propagated from T2 transgenic lines, which were confirmed to carry the recombinant pEG101 vector and displayed enhanced transcript and protein expression levels of the respective transgene (detailed in section 4.3.3). The lines also exhibited 100% survival on BASTA MS selection plates further discussed in section 4.4.1, verifying their homozygosity. The two AT60 overexpressing lines had over double the AT60 expression levels relative to Col-0, with considerably higher levels in 60B-5, validating the lines were overexpressing the respective transgene to a significant level. AT80 transcript levels also considerably increased in comparison to Col-0, which was consistent with the trend observed in the T2 transgenic line (60A in Figure 4.3-A). For the AT80 overexpressing lines, transcript levels of AT80 were

significantly greater than WT, with considerably higher levels observed in 80D-6. Therefore, both lines could be considered genuine AT80 gain-of-function mutants. As observed in the T2 parent AT80 overexpressing lines (80C and 80D in Figure 4.3A), heightened AT80 expression levels caused a significant increase in AT60 expression levels relative to Col-0. Therefore, this real-time PCR experiment confirmed the four T3 AT60 and AT80 lines as overexpressing plants and were subsequently subjected to various phenotyping analyses.

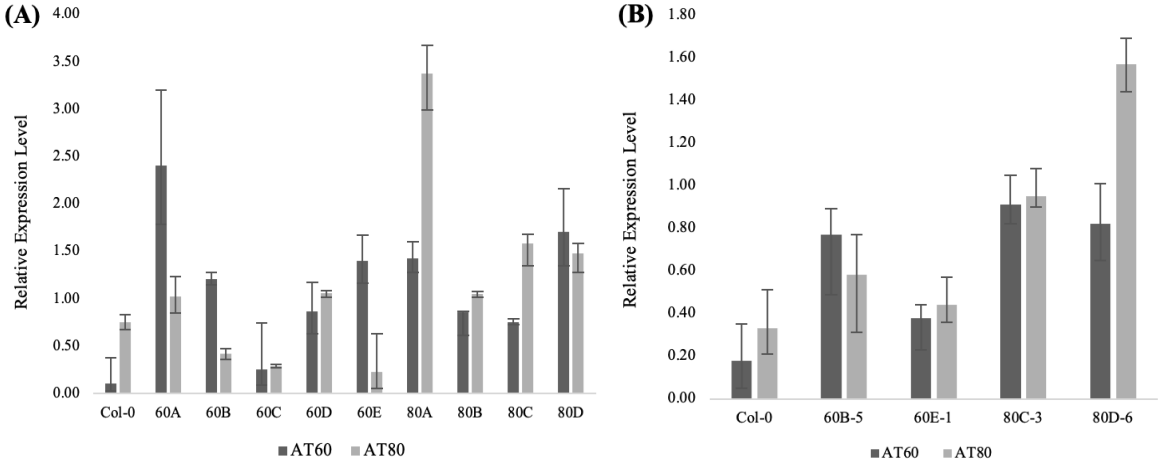


Figure 4.3 – Expression levels of AT60 and AT80 measured using real-time PCR of Col-0 and T2 and T3 AT60 and AT80 overexpressing plants. RNA was extracted from three-week old Col-0 Arabidopsis plants and both T2 and T3 AT60 and AT80 overexpressing plants. cDNA was synthesised from these RNA samples and using real-time PCR, their AT60 and AT80 transcript levels were measured and normalised against the expression of actin in Arabidopsis. A) Displays the qPCR results from T2 AT60 and AT80 overexpressing transgenic lines. B) Displays the qPCR results from T3 AT60 and AT80 overexpressing lines. Data presented are means ± SE of technical repeats conducted in triplicates, however, due to time constraints, the experiment was only repeated once.

4.3.3. Immunoblot analyses of T2 overexpressing transgenic lines

To confirm the presence and increased expression of the YFP-tagged AT60 or AT80 recombinant protein in the T2 overexpressing lines, the total protein was extracted from the gain-of-function mutants and analysed by SDS-page and immunoblotting. Seedlings of T2 overexpressing transgenic lines were grown for 10 days in normal conditions and subsequently frozen in liquid nitrogen. The total protein was then extracted from each T2 overexpressing transgenic line. Control samples were also analysed. This comprised of 10-day old Col-0 Arabidopsis seedlings which served as the negative YFP control, and samples from the transient expression assay. This included YFP-only infiltrated *N. benthamiana* samples acting as the positive YFP control, and YFP:HA:AT60 and YFP:HA:AT80 infiltrated *N. benthamiana* samples functioning as the positive control for the respective recombinant YFP- and HA-tagged DeSi proteins. The concentration of all protein samples was equilibrated before being separated on a SDS-PAGE gel and antibodies raised against YFP were used to probe the immunoblot to detect protein presence.

Figure 4.4 displays the western blot results of the total extracted protein from all analysed samples split by AT60 (A) and AT80 (B). For both the AT60 and AT80 representative blots, the first lane contained the protein extracted from *N. benthamiana* leaves infiltrated with the pEG101 vector only and displayed the expected band

for the YFP protein at approximately 27kDa. The second lane with Col-0 extracted protein showed no protein band as the native protein (AT60 or AT80) was not fused to YFP. The next lane in Figure 4.4-A and -B displayed the protein extracted from *N. benthamiana* leaves infiltrated with AT60 and AT80, respectively, cloned into the pEG101 vector. The YFP- and HA-tagged AT60 (Figure 4.4-A) and AT80 (Figure 4.4-B) recombinant protein samples had an expected molecular weight of 58kDa and 57kDa, respectively, serving as a positive control sample for the T2 overexpressing transgenic Arabidopsis samples.

For the AT60 overexpressing samples displayed in Figure 4.4-A, the YFP- and HA-tagged AT60 recombinant protein with an expected molecular weight of 58kDa could be observed in all lines (60A - 60E). This confirmed the recombinant protein was expressed in all T2 AT60 gain-of-function mutant lines. However, the protein was evidently expressed at a higher level in the 60A, 60B and 60E line, which was consistent with the qPCR results. For the AT80 overexpressing samples, all lines except for 80A displayed the band corresponding to the YFP- and HA-tagged AT80 recombinant protein at the expected molecular weight of 57kDa. Therefore, this western blot analysis confirmed the presence and elevated protein level of AT80 in the 80B, 80C and 80D lines. Although greater expression could be observed in the 80B line, this result was inconsistent with the qPCR results which found only a slight increase in AT80 transcript levels.

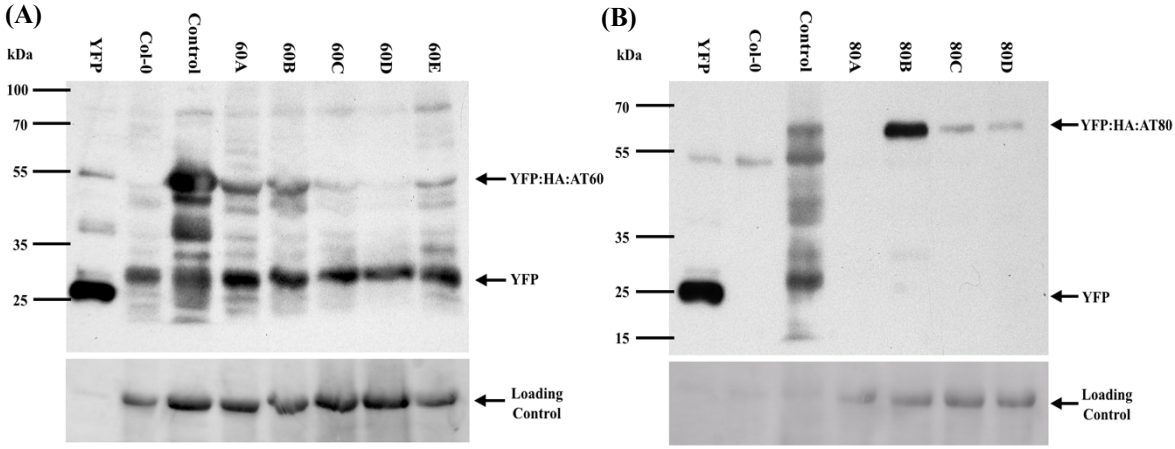


Figure 4.4 – Western blot to identify the expression of the recombinant AT60 and AT80 protein in T2 overexpressing lines. Total protein was extracted from ten-day old Col-0 and T2 AT60 and AT80 overexpressing transgenic seedlings. Samples were separated on a SDS-PAGE gel and protein bands were visualised with a western blot using antibodies raised against YFP. Samples from 10-day old Col-0 seedlings (Col-0), YFP-only infiltrated *N. benthamiana* (YFP) and YFP:HA:AT60 and YFP:HA:AT80 infiltrated *N. benthamiana* (Control) were run alongside the overexpressing transgenics samples as controls. A) Displays the western blot results for T2 AT60 overexpressing lines (60A-60E) with respective control samples. B) Presents the western blot results for T2 AT80 overexpressing transgenics lines (80A-80D) with respective control samples. The top panel displays the bands at the expected size for the YFP protein (27kDa) and recombinant YFP:HA:AT60 (58kDa) or YFP:HA:AT80 (57kDa) as indicated by the arrow. The bottom panel presents the total extract as a loading control for the normalisation of the samples.

The genotyping, protein and expression analysis experiments were conducted to determine which T2 overexpressing lines should be propagated to the T3 generation. The three experiments deduced that for the

AT60 overexpressing lines, 60B and 60E lines were selected, and for AT80 overexpressing lines, 80C and 80D lines were chosen for propagation. The RT-PCR verified the four selected transgenic lines were carrying the recombinant pEG101 vector (Figure 4.2). The real-time PCR results established there was an increased expression of AT60 levels in 60E and 60B, and AT80 levels in 80C and 80D lines (Figure 4.3-A), relative to WT. Finally, the western blot determined the presence and elevated expression of the respective YFP-tagged recombinant DeSi protease in the four selected overexpressing lines (Figure 4.4). Therefore, these T2 overexpressing lines were selected for T3 propagation and subsequent phenotyping analyses.

4.4. Phenotypic Characterisation of the AT60 and AT80 Overexpressing Lines

4.4.1. Screening T3 Overexpressing Transgenic Plants

The selected T2 AT60 and AT80 overexpressing lines were propagated to T3 generation and the seeds from these T2 lines were again screened on BASTA (30µg mL^{-1}) MS selection plates for 100% survival. The T3 lines with 100% resistant seedlings presented in Table 4.1, could be considered homozygous overexpressing plants. From the 100% surviving T3 overexpressing lines in Table 4.1, a real-time PCR was conducted on all lines with results displayed in Figure 4.3-B. The T3 overexpressing transgenic lines, 60B-5, 60E-1, 80C-3 and 80D-6 exhibited significantly enhanced expression of the respective DeSi protease. Therefore, seedlings from these transgenic lines were subsequently subjected to phenotypic analyses.

Table 4.1 – T3 transgenic lines with 100% survival on BASTA (30µg mL^{-1}) MS selection plates from the segregation analysis

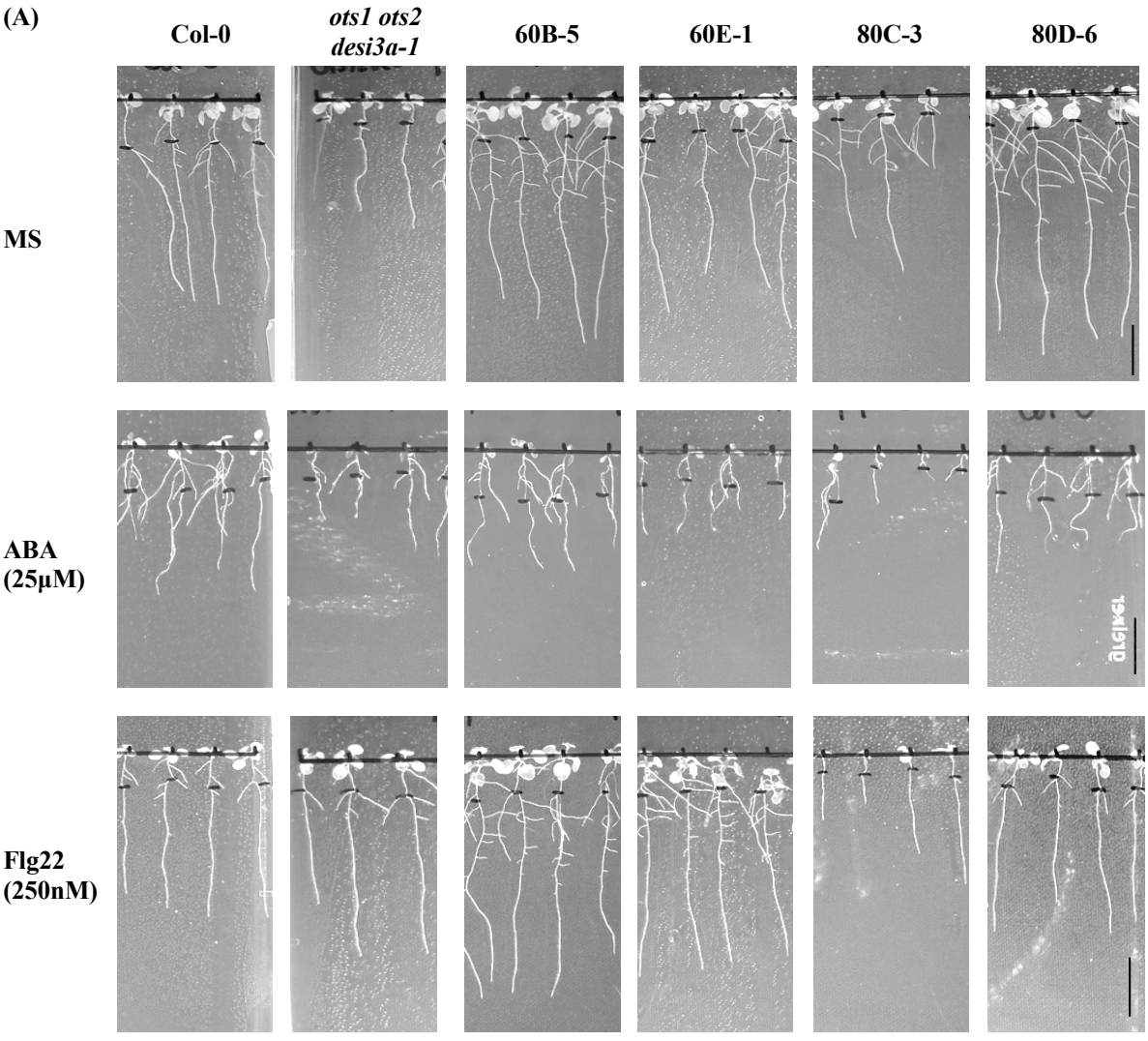
Name	Construct	Background
60A-4	35S:60B-YFP-HA	Col-0
60B-1	35S:60E-YFP-HA	Col-0
80A-1	35S:80C-YFP-HA	Col-0
80A-4	35S:80C-YFP-HA	Col-0
80A-6	35S:80C-YFP-HA	Col-0
80B-1	35S:80D-YFP-HA	Col-0
80B-2	35S:80D-YFP-HA	Col-0
80B-5	35S:80D-YFP-HA	Col-0

4.4.2. Root Length and Fresh Weight Assay on T3 Overexpressing Plants

To phenotype the selected overexpressing transgenics from the T3 60B-5, 60E-1, 80C-3 and 80D-6 lines, a root length and fresh weight assay in normal and stress conditions was performed. Both assays were performed following the exact same method used to analyse the AT60-AT80KO mutant lines. These assays assess whether the T3 overexpressing plants display pleiotropic phenotypes in root growth and biomass under normal conditions as well as in response to ABA and flg22. This experiment would also deduce whether the enhanced expression of the DeSi proteases would result in a mutant phenotype opposite to the phenotype observed in the analysis of the KO transgenics. Therefore, both assays on the single AT60 and AT80 overexpressing plants will help further characterise the role of the DeSi proteases in Arabidopsis development and stress-induced responses.

4.4.2.1. Root Length Assay on T3 Overexpressing Transgenic Plants

For the root length assay, Col-0, *ots1 ots2*, *desi3a-1* and the T3 single AT60 and AT80 overexpressing transgenic lines were tested. The *ots1 ots2* and *desi3a-1* lines were used as comparative samples as both are KO mutants of Arabidopsis SUMO proteases, especially as the latter is a DeSi SUMO protease mutant. All seedlings were grown on normal MS media for 4 days and then transferred to MS media, ABA (25µM) MS media to simulate abiotic stresses, or flg22 (250nM) MS media as a biotic stress inducer. 6 days after the transfer, the root lengths of all seedlings were measured digitally using the Fiji software. Representative samples for each genotype on the three mediums are displayed in Figure 4.5-A. Quantification of the root lengths were calculated using averages of all seedlings analysed per genotype. At least 15 individual seedlings for each genotype were analysed and the experiment was only repeated once due to time constraints. The average root length for each genotype grown on MS media, ABA- and flg22-treated MS media are presented in Figure 4.5-B. The Kruskal-Wallis test followed by a Bonferroni post hoc test was used to calculate statistical significance with the results displayed in Figure 4.5-B. To quantify the effect the stress inducers had on root elongation for all genotypes, the difference between the average root length of seedlings grown on MS media and the averages of seedlings grown on ABA- or flg22-supplemented MS media of the same genotype was calculated. This assay would identify whether the difference in root length was caused by overall slower growth of the genotype or the presence of ABA or flg22 hindering root elongation in the seedlings. Results of this assay are displayed in Figure 4.5-C and statistical significance was calculated using the Mann-Whitney U-test.



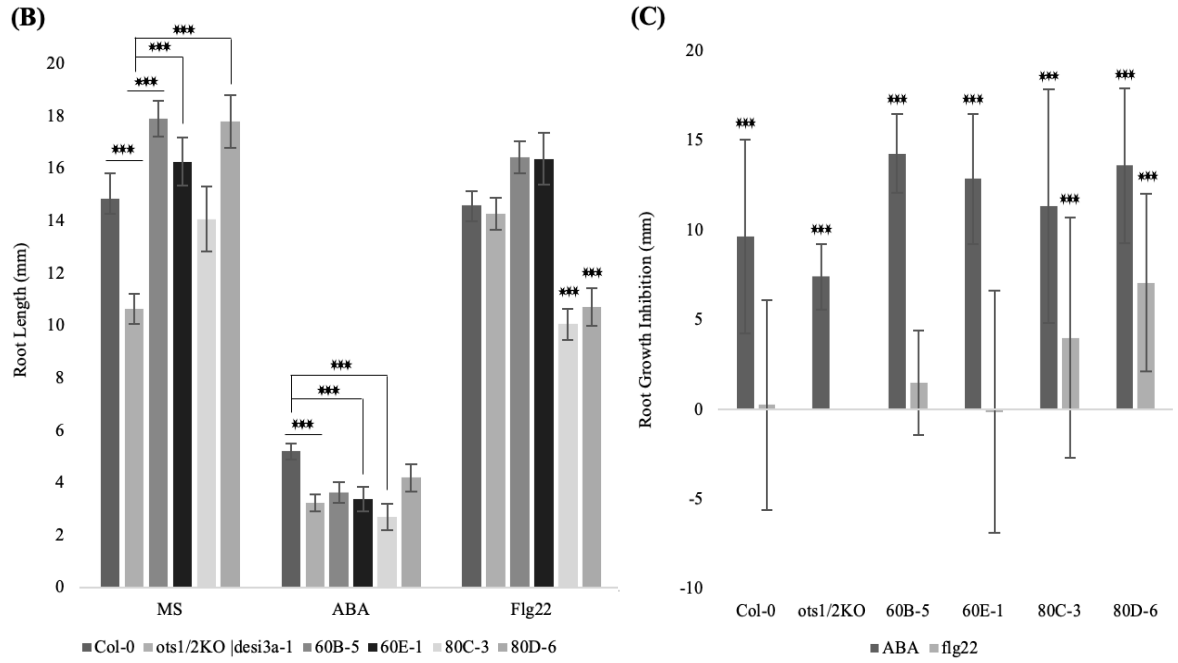


Figure 4.5 – Results from the Root Length Assay. A) Phenotypic appearance of representative 10-day old seedlings from each genotype grown on normal MS plates and ABA (25μM) and flg22 (250nm) containing MS plates for 6 days. Scale bar equals 1cm. B) Quantification of average root growth for each genotype grown on normal and treatment-containing media. C) Quantification of average root growth inhibition for each genotype comparing the difference between root lengths of seedlings grown in normal and stress-induced conditions (ABA and flg22). Data presented are means ± SE from at least 15 individual seedlings for each genotype. Significance was assessed using the Mann-Whitney U test for comparing between 2 samples (C) and Kruskal-Wallis test followed by Bonferroni post-test for comparing between 3 samples or more (B). Significance values: * = $P < 0.05$; ** = $P < 0.01$; *** = $P < 0.005$.

Looking at Figure 4.5-B, the Kruskal-Wallis test determined a statistically significant difference in mean root length between genotypes grown on all 3 mediums. Post-hoc tests revealed that in normal conditions, the root length of *ots1 and ots2* was significantly shorter than all other lines apart from 80C-3. All other genotypes were found to have equal mean root lengths. In ABA-induced conditions, Col-0 had the longest root elongation out of all genotypes, however, the root length of Col-0 was only significantly longer than *ots1 and ots2*, 60E-1 and 80C-3, with no significant difference to 60B-5 and 80D-6 lines. When comparing average root lengths between genotypes of seedlings grown on flg22 MS media, the root length of 80C-3 and 80D-6 were significantly shorter than all other genotypes. The two AT60 overexpressing lines, 60B-5 and 60E-1, exhibited longer root lengths than Col-0 and *ots1 and ots2*, however not to a significant extent. For the root growth inhibition assay displayed in Figure 4.5-C, the Mann-Whitney U test found that for all genotypes, the presence of ABA significantly inhibited root growth. ABA inhibited the root growth of the overexpressing lines the most and *ots1 and ots2* the least, implying the two DeSi proteases may be involved in the abiotic stress response pathway. Exogenous application of flg22 only significantly inhibited the root elongation of the two AT80 overexpressing transgenic lines, 80C-3 and 80D-6, indicating AT80 could be implicated in the flg22 response pathway.

4.4.2.2. Fresh Weight Assay on T3 Overexpressing Transgenic Plants

Fresh weight data was collected from 10-day old seedlings to further investigate the phenotypic differences between the T3 overexpressing transgenic lines and control lines. This assay also determined whether the phenotypes observed in the root length assay were a result of the stress inducer inhibiting root growth mechanisms or impeding overall growth of all plant tissues. The seedlings grown on normal MS and ABA- and flg22-containing MS for 6 days were weighed and the average of all weighed seedlings for each genotype was calculated. At least 15 individual seedlings for each genotype was analysed and the experiment was only repeated once due to time constraints. To coherently present the statistical analysis results, the graphs were divided and displayed by their comparative groups in Figure 4.6.

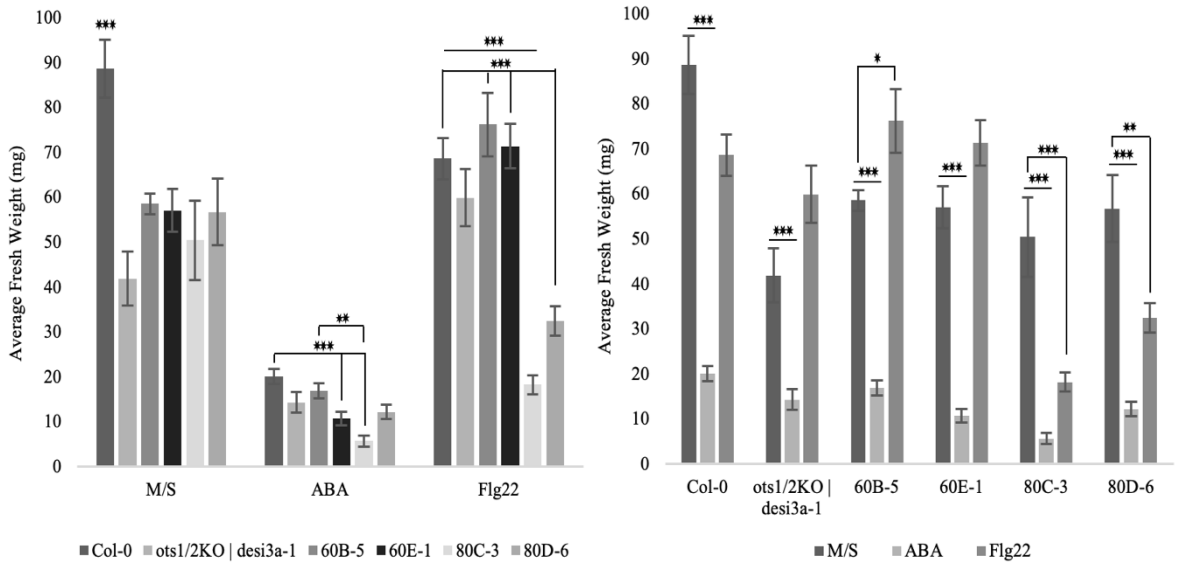


Figure 4.6 – Results from the Fresh Weight Assay. The figure displays the quantification of the biomass of 10-day old seedlings exposed to treatment and no treatment for 6 days. Seedlings were grown on MS media for 4 days then transferred to plates with normal MS media and MS media supplemented with ABA (25µM) and flg22 (250nm). Following 6 days of further growth, the plants were carefully removed and weighed. The left graph compares the average fresh weight of seedlings between genotypes for each medium. The right graph compares the effect of ABA and flg22 on each genotype. Data presented are means ± SE from at least 15 individual seedlings for each genotype. Significance was assessed using Mann-Whitney U test for comparing 2 samples (left) and Kruskal-Wallis test followed by Bonferroni post-test for comparing 3 samples or more (right). Significance values: * = $P < 0.05$; ** = $P < 0.01$; *** = $P < 0.005$.

The Kruskal-Wallis test determined there were statistically significant differences in the biomass between the genotypes of seedlings grown on all three mediums. In normal conditions, the Bonferroni post-hoc test revealed Col-0 weighed significantly more than all other transgenic lines. In ABA conditions, aligning with the root length assay, Col-0 was again significantly heavier than 60E-1 and 80C-3, however, no difference in biomass was reported between WT and the 60B-5 and 80D-6 lines. There was also a significant variation in biomass between 60B-5 and 80C-3. Exogenous flg22 application caused a significant reduction in the fresh weight of both AT80 overexpressing lines in comparison to all other genotypes. These findings were consistent with the root length assay indicating that for AT60 and AT80 overexpressing transgenics, the presence of ABA and flg22 could either effect or have no effect on the overall growth of all plant tissues including root development.

When comparing the biomass of seedlings grown on various mediums for each genotype, the seedlings grown on ABA mediums were significantly lighter than seedlings grown on MS media for all genotypes. The reduction in fresh weight of ABA-exposed seedlings to normally grown was approximately 77%, 66%, 71%, 81%, 89% and 79% for Col-0, *ots1 ots2*, 60B-5, 60E-1, 80C-3 and 80D-6, respectively. These results were consistent with the root length assay implying that across all genotypes, exogenous ABA triggers a growth arrest in all plant tissues including the inhibition of root growth. Consistent with the root length assay, the presence of flg22 significantly reduced the weight of 80C-3 and 80D-6 lines, thereby highlighting a possible role of AT80 in modulating immune signalling to trigger the switch from growth to defence mode in Arabidopsis plants. Interestingly, 60B-5 lines in flg22-induced conditions were significantly heavier than seedlings grown in normal conditions. The root length and fresh weight assay findings indicate AT60 and AT80 may have a role in modulating abiotic stress and immune signalling in Arabidopsis plants.

4.4.3. Subcellular Localisation Studies of the YFP-tagged DeSi Proteases

To date, there is limited data available regarding the subcellular localisation of DeSi SUMO proteases. Very recently, a study was able to identify the Desi3a protein to be localised in the plasma membrane (Orosa *et al.*, 2018). To further investigate the function of the two DeSi proteases, subcellular localisation studies were conducted on the T3 AT60 and AT80 overexpressing lines. The stable transformation and expression of the pEG101 constructs in Arabidopsis plants would provide more insight into the protein's subcellular localisation in comparison to the transient expression assay conducted in *N. benthamiana* plants. This subcellular localisation study was first performed in normal conditions to gain a better understanding of the organ-specific and cellular localisation of the two DeSi proteases. Subsequently, the study was repeated in stress-induced conditions for comparative measures using the phytohormone ABA and the pathogen response elicitor flg22. The YFP-tagged proteins were visualised using the Zeiss LSM 880 microscope. An empty vector expressing only YFP was used as a control to verify the tag was not guiding the localisation of the proteases, as well as Col-0 samples as a negative control.

4.4.3.1. Subcellular localisation of YFP-tagged DeSi proteases in T3 AT60 and AT80 overexpressing lines in normal conditions

The subcellular localisation of the YFP-tagged DeSi protease in the T3 AT60 and AT80 overexpressing lines were analysed to determine the spatial expression patterns of both recombinant proteins. However, as the expression of the proteases are driven by a non-native promoter, the experiment results should not be considered reflective of the endogenous sub-cellular localisation of AT60 and AT80. Seedlings from the transgenic lines 60B-5, 60E-1, 80C-3 and 80D-6 were grown on BASTA (30µg mL^{-1}) MS selection plates for 4 days before being mounted onto a microscope slide and visualised with the Zeiss LSM 880 microscope. Expression levels of the YFP-tagged DeSi protease were examined in the leaf epidermis, stem, root and root tip organs of the seedlings. In addition to the T3 AT60 and AT80 overexpressing transgenic lines, Col-0 and YFP control seedlings were also examined. Approximately 3 individual seedlings from each line were tested and representative images of the analysed seedlings per genotype are presented in Figure 4.7. Samples of Col-0 seedlings displaying no YFP fluorescence are displayed in Appendix Figure 8.8. Total mean fluorescence of

the YFP-tagged proteins was measured using the Fiji software to determine the variation in protein expression levels across all analysed samples. The total mean fluorescence of the YFP-tagged DeSi proteases for each seedling across the various plant tissues are presented in Figure 4.8.

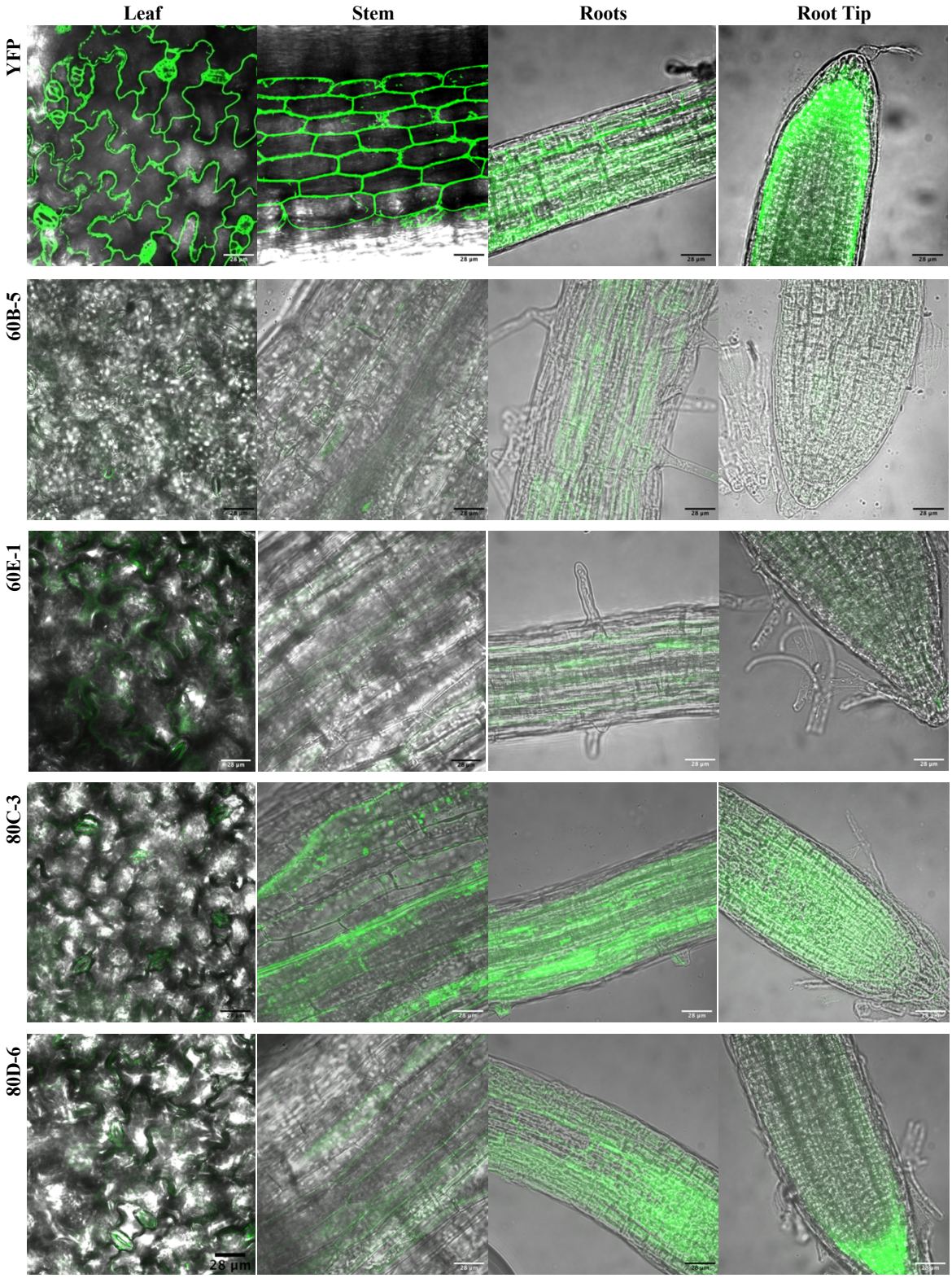


Figure 4.7 – Recombinant AT60 and AT80 subcellular localisation in the T3 overexpressing transgenic lines. The recombinant pEG101 construct overexpressing either the AT60 or AT80 protein was transformed into Arabidopsis plants and propagated to T3 generation. Seedlings from the T3 AT60 and AT80 overexpressing transgenic lines were initially grown on BASTA (30µg mL^{-1}) MS selection plates before being mounted on microscope slides and viewed by a Zeiss LSM 880 microscope. The vertical panels in the figure denote the plant tissue of the seedling examined and the horizontal panels indicate the genotype of the analysed seedling. The "YFP" panel displays the localisation of the YFP protein from the YFP-only transformed Arabidopsis seedlings. Scale bar = 28 µm.

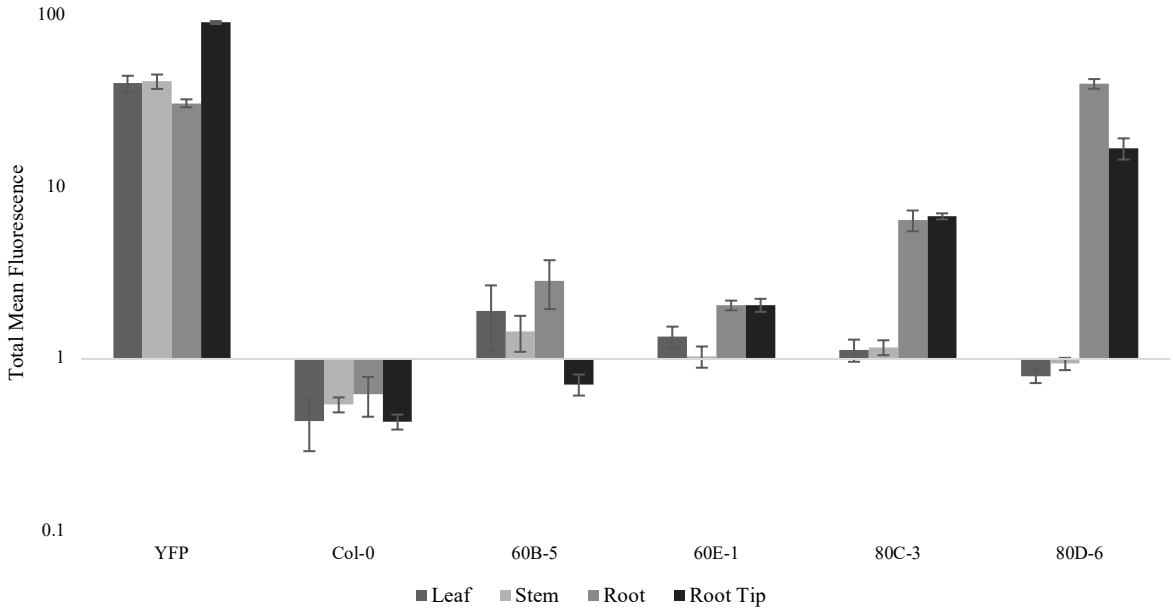


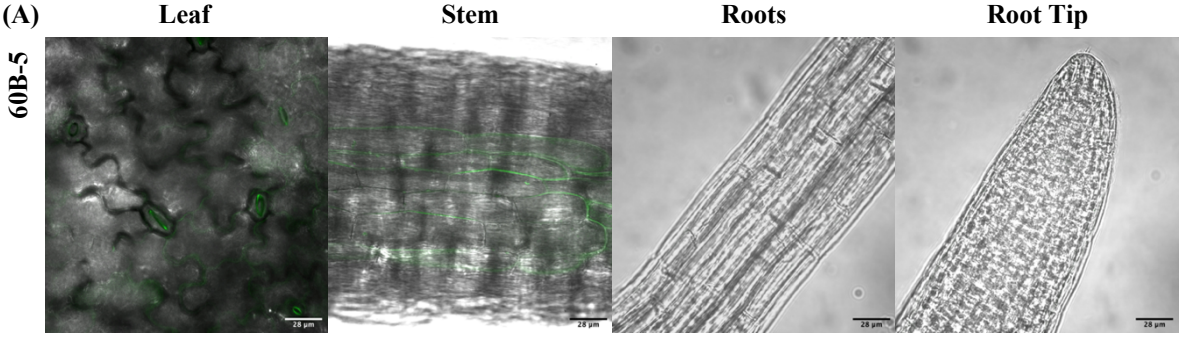
Figure 4.8 – Total mean fluorescence of the YFP-tagged proteins in each plant tissue for all analysed genotypes. Fluorescence of the YFP-tagged proteins were measured using the Fiji software. The mean fluorescence of the AT60 protein was measured in the T3 AT60 overexpressing seedlings (60B-5 and 60E-1) and the AT80 protein was measured in the AT80 overexpressing seedlings (80C-3 and 80D-6). Measurements were taken on over 5 images of each section from at least 3 different individual seedlings per genotype; $n \geq 14$. Data presented are means ± SE and the Y-axis is on a logarithmic scale.

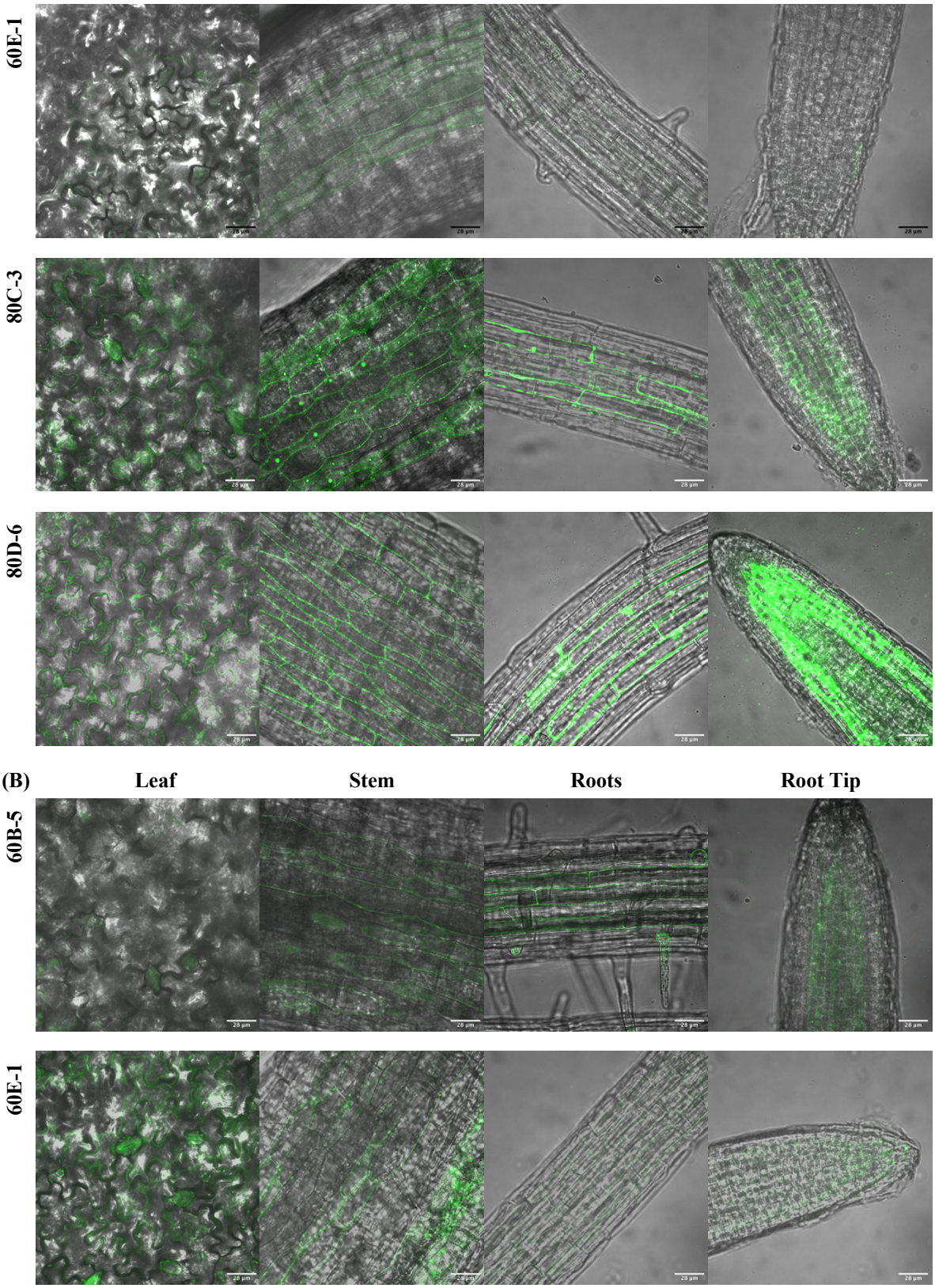
In Figure 4.7, as expected, the YFP protein in the YFP-only transformed Arabidopsis seedlings, localised across all subcellular structures and plant organs. The mean fluorescence of YFP was very high relative to all other compared samples across all four analysed tissues. On the contrary, the mean fluorescence of Col-0 seedlings across all tissues was as expected, significantly low and close to zero as shown in Figure 4.8. There was low overall expression of the AT60 protein in the T3 AT60 overexpressing lines, 60B-5 and 60E-1. This was reflected in the mean fluorescence of the YFP-tagged AT60 protein in both lines across all plant tissues. Although very faintly, AT60 could be seen to localise outside the nucleus towards the cell periphery, most likely the plasma membrane, in all four analysed tissues. However, with very low protein expression and difficulty in distinguishing the recombinant protein from autofluorescence, this could not be conclusive. For the T3 AT80 overexpressing transgenic lines, 80C-3 and 80D-6, there was greater overall expression of the AT80 protein. Across all four examined tissues, the AT80 protein localised to the cell periphery, most likely the plasma membrane. In the leaf epidermal cells, the AT80 protease seemed to also localise to the guard cells.

However, with relatively low fluorescence in the leaf, this could be autofluorescence. Alternatively, the mean fluorescence of AT80 in the root and root tip cells in both 80C-3 and 80D-6 lines were relatively high. Therefore, this experiment determined the AT80 DeSi protease accumulates at high concentrations in the root and root tip cells of Arabidopsis plants, localising towards the cell periphery and plasma membrane. The findings that the two DeSi proteases localised outside the nucleus towards the cell periphery were to an extent, consistent with the *N. benthamiana* transient expression assay described in section 3.2.3 in the previous chapter. However, as the expression of the proteases were driven by a non-native promoter, the findings from this experiment do not accurately reflect the endogenous sub-cellular localisation of AT60 and AT80.

4.4.3.2. Subcellular localisation of YFP-tagged DeSi proteases in T3 AT60 and AT80 overexpressing transgenic lines in response to ABA and flg22

The subcellular localisation of the YFP-tagged DeSi proteases were analysed to determine the spatial expression patterns of both AT60 and AT80 proteins in response to stress inducers. This experiment would supplement the results observed in the qPCR experiment and stress response phenotyping assays. Seedlings from the T3 AT60 and AT80 overexpressing lines were grown on BASTA (30µg mL^{-1}) MS selection plates for 4 days before being transferred into a Petri dish containing liquid MS media. 10µM of ABA and 200nM of flg22 was added to the liquid MS media and left on a shaking rocker for 2 hours for ABA and 30 minutes for flg22. The seedlings were then immediately transferred onto a microscope slide and visualised with the Zeiss LSM 880 microscope. Expression levels of the YFP-tagged DeSi proteases were examined in the leaf, stem, root and root tips organs of the seedlings. This experiment was also performed on YFP control seedlings to confirm the treatments of ABA and flg22 had no influence on the YFP protein. Samples of the treated YFP control seedlings displaying no difference in spatial expression patterns or levels are displayed in the Appendix Figure 8.8. Approximately 3-5 individual seedlings from each genotype was analysed and representative images of the samples from each examined tissue in response to ABA and flg22 are displayed in Figure 4.9-A and Figure 4.9-B, respectively. Total mean fluorescence of the YFP-tagged proteins was measured using the Fiji software to determine whether the expression level of the DeSi proteases significantly differed in response to stress inducers. The total mean fluorescence of the YFP-tagged AT60 and AT80 protein for all seedlings exposed to ABA and flg22 treatments per genotype are presented in Figure 4.10. In both figures, the protein fluorescence in stress-induced conditions was compared to the protein fluorescence in normal conditions and statistical significance was calculated using the Mann Whitney U-test. The YFP protein in YFP control seedlings displayed no difference in spatial expression pattern or total mean fluorescence when exposed to ABA or flg22 as depicted in Appendix Figure 8.8 and Figure 8.9.




60E-1
80C-3
80D-6
(B)
Leaf
Stem
Roots
Root Tip
60B-5
60E-1

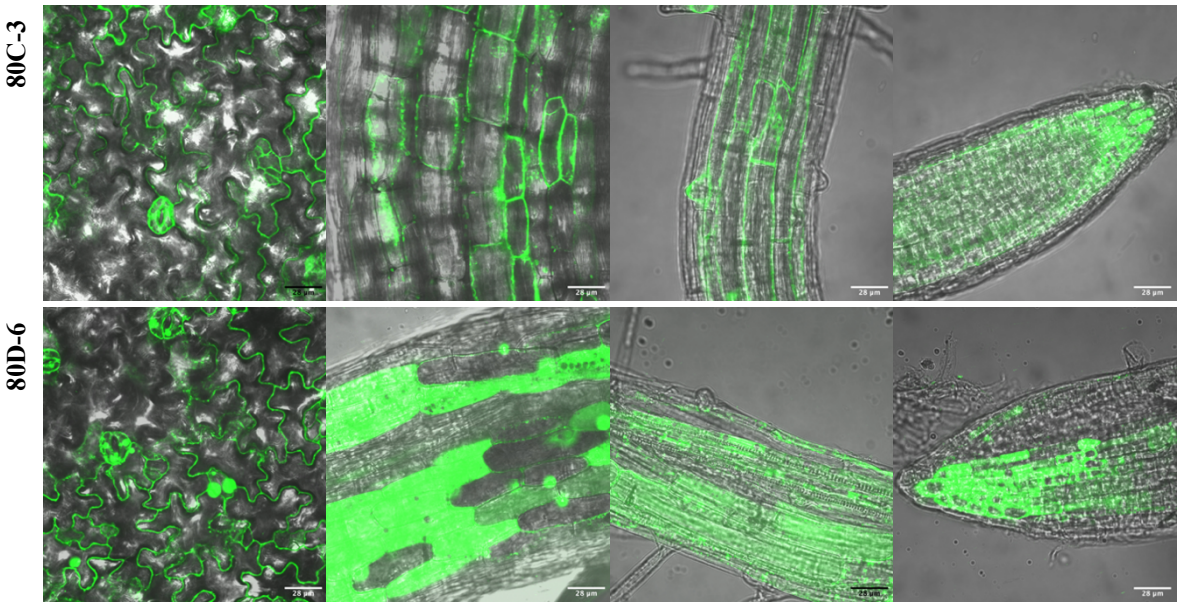
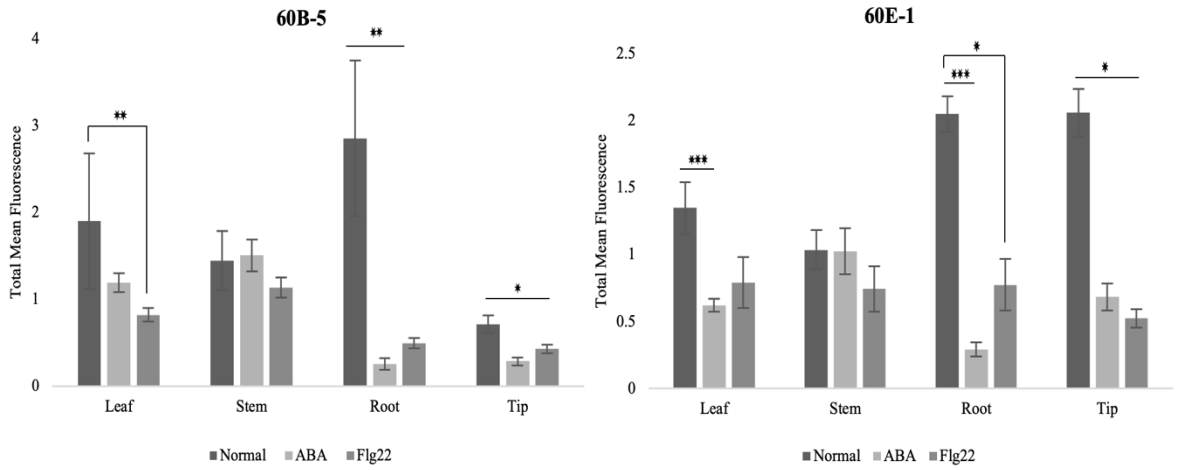


Figure 4.9 – Subcellular localisation of AT60 and AT80 in response to ABA and flg22 in T3 overexpressing lines. Seeds from the T3 AT60 and AT80 overexpressing lines were first grown on BASTA (30µg mL^{-1}) MS selection plates for 4 days. Subsequently, the seedlings were transferred into ABA (10µM) or flg22 (200nM) supplemented liquid MS media for 2 hours and 30 minutes, respectively. The seedlings were then mounted on microscope slides and viewed by a Zeiss LSM 880 microscope. A) Displays the representative image of seedlings per genotype exposed to ABA (10µM) for 2 hours. B) Presents the representative image of seedlings per genotype exposed to flg22 (200nM) for 30 minutes. Scale bar = 28 µm.



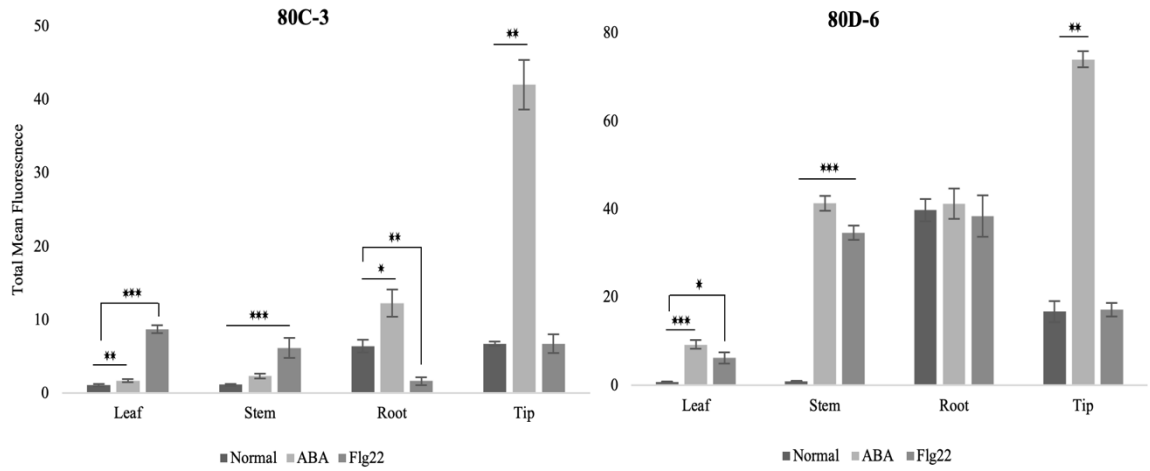


Figure 4.10 – Total mean fluorescence of the YFP-tagged proteins in the T3 AT60 and AT80 overexpressing seedlings under normal and stress-induced conditions. Fluorescence of the YFP-tagged proteins in each condition was measured using the Fiji software. In the AT60 overexpressing lines (60B-5 and 60E-1) on the top row, the mean fluorescence of the AT60 protein in response to ABA (10μM) and flg22 (200nM) was measured and compared against AT60 fluorescence in normal conditions. In the AT80 overexpressing lines (80C-3 and 80D-6) on the bottom row, the average fluorescence of the AT80 protein when exposed to ABA (10μM) and flg22 (200nM) was measured and compared to the normal average fluorescence of the DeSi protease. Measurements were taken from over 5 images of each plant organ from at least 3 different individual seedlings per genotype for each treatment; $n \geq 14$. Data presented are means ± SE and significance between average fluorescence of the protein in normal and ABA- or flg22-induced conditions were assessed using the Mann-Whitney U test. Significance values: * = $P < 0.05$; ** = $P < 0.01$; *** = $P < 0.005$.

In ABA-induced conditions (Figure 4.9-A), the AT60 protein in both T3 AT60 overexpressing lines (60B-5 and 60E-1) seemed to have very little expression across the four plant tissues. Similar to the findings in normal conditions, the AT60 protease appeared to be localised to the plasma membrane in both leaf and stem tissues, though very faintly. No AT60 protein could be detected in the root or root tip cells. This was reflected in the total mean fluorescence results of AT60 (Figure 4.10), where a significant reduction in expression levels was observed in the root and root tip cells relative to normal conditions. Within the leaf cells, AT60 protein fluorescence also decreased in response to ABA, however, this was only significant in one of the transgenic lines (60E-1). In the stem tissue of the seedlings, no notable difference in AT60 fluorescence was observed when exposed to ABA. This experiment confirmed that in response to ABA, the expression of the AT60 DeSi protease considerably decreased in the root system and leaves of Arabidopsis plants, however, no major difference was detected in the subcellular localisation of AT60. On the other hand, there was a notable increase in the expression of the AT80 DeSi protease in the T3 AT80 overexpressing seedlings (80C-3 and 80D-6) exposed to ABA treatment (Figure 4.9-A). The subcellular localisation of AT80 remained unchanged with the localisation observed in the normal conditions, with the protease localising outside the nucleus towards the plasma membrane. However, the strength of the fluorescence was significantly higher in the ABA-exposed transgenic seedlings relative to normal conditions across all plant tissues. The AT80 protein fluorescence was significantly greater in the root tip cells of the ABA-exposed seedlings relative to normally-grown seedlings for both tested lines. This experiment deduced that in response to ABA treatment, the expression of the AT80 DeSi protease significantly increased across nearly all plant tissues of the analysed seedlings.

When the AT60 overexpressing lines were exposed to flg22, similar expression results to the ABA-exposed seedlings were observed. AT60 expression significantly decreased in the root and root tip cells for both overexpressing lines (Figure 4.9-B), whilst in the leaf cells, only reducing by a significant amount in one overexpressing line (60B-5). No difference in AT60 fluorescence or spatial expression pattern could be detected in the stem tissue of seedlings treated with flg22 than in normal and ABA-induced conditions. In the leaf epidermal cells, AT60 seemed to localise to the guard cells as well as the plasma membrane, however, the decrease in AT60 fluorescence level indicated this could be chlorophyll autofluorescence and not the recombinant protein. The findings conclude that in response to flg22, AT60 protein abundance significantly decreased across all plant tissues excluding the stem, however, AT60 subcellular localisation remained the same. In response to flg22, there was strong overall expression of AT80 in the AT80 overexpressing seedlings. The AT80 protein localised to the same subcellular structure of the seedling as observed in normal conditions. The presence of the AT80 DeSi protease could be detected in the plasma membrane across all plant organs. Within the leaf epidermal cells, the AT80 protein also localised to the guard cells as can be observed in Figure 4.9-B. The fluorescence intensity of AT80 was significantly greater in the transgenic seedlings exposed to flg22 in the leaf epidermal and plant stem tissue. Interestingly, AT80 protein fluorescence either remained the same or significantly reduced in the root and root tip cells of flg22-treated seedlings relative to normally grown seedlings. This experiment established that in response to pathogen infection, the AT80 DeSi protease remained localised to the plasma membrane, however, protein abundance reduced in the roots whilst concentrating in the leaf epidermal and stem cells of Arabidopsis plants.

4.5. Discussion

This chapter explores how the single gene overexpressing transgenic plants were generated and the subsequent genotyping analyses performed on the transgenic lines to confirm their gain-of-function attributes. Single overexpressing transgenic plants were created and analysed to help deduce whether AT60 and AT80 DeSi proteases function redundantly or divergently as well as for functional characterisation. Various phenotyping experiments were conducted on the confirmed homozygous single gene overexpressing lines. These analyses comprised of the root length and fresh weight assay in response to stress elicitors, as well as confocal microscopy experiments to analyse the subcellular localisation of the overexpressed AT60 and AT80 protease in normal and stress-induced conditions. The phenotyping analyses would further corroborate whether AT60 and AT80 are implicated in the development and stress response pathway of Arabidopsis plants.

4.5.1. Genotyping experiments validated the transgenic lines as AT60 and AT80 homozygous overexpressing transgenic lines

Recombinant *Agrobacterium* cells overexpressing AT60 and AT80 in the pEG101 vector were used to generate the single gene overexpressing Arabidopsis plants. The floral-dipping method was used to transform the Col-0 plants with the recombinant vectors. T1 seedlings were screened for positive transformants on BASTA-treated soil and grown for seed harvesting. Resistant seedlings from T1 lines segregating at 3:1, resistant to non-resistant seedlings, on BASTA selection plates, were grown in normal conditions for genotyping and

immunoblot analyses. The exact same protocol as Conti *et al.* (2008) was used to generate and analyse overexpressing transgenic lines. In the published study, the researchers produced a single overexpressing OTS1 transgenic lines using the pEG104 vector.

To confirm the T2 overexpressing plants were truly gain-of-function mutants, a RT-PCR, real-time PCR and immunoblot analysis was performed. These genotyping analyses were all conducted at the T2 generation, where the lines were already confirmed to carry single copies of the transgene insertion as they segregated at a 3:1 ratio following Mendelian law of segregation (Konstantinova *et al.*, 2003). Therefore, the verified T2 gain-of-function mutant lines would generate homozygous T3 overexpressing seedlings segregating at 100%. The methods used to genotype and analyse the overexpressing transgenic plants were consistent with published papers (Conti *et al.*, 2008; Lin *et al.*, 2016). However, some studies that also generated overexpressing plants, conducted a genomic Southern blot analysis in addition to performing a RT-PCR to validate transgene insertion. The Southern blot analysis was conducted to determine the copy number of transgene insertions, thereby confirming a stable transformation event had occurred (Kim *et al.*, 2017). If this experiment were repeated, a Southern blot analysis would have been performed on the overexpressing transgenic lines for stable transformation validation.

The RT-PCR identified which T2 transgenic lines were successfully transformed and carrying the recombinant pEG101 vector. The presence of the YFP coding sequence was verified in the 60B, 60E, 80C and 80D transgenic lines, which were subsequently subjected to further genotyping analyses. However, multiple bands could be observed in the actin controls, which may be a consequence of off-target gene amplification. Additional PCRs would need to be carried out at a higher temperature to ensure no off-target genes were amplified. With non-specific bands appearing in the RT-PCR, higher annealing temperatures were used in the following real-time PCR experiment which used actin as the housekeeping gene, in order to prevent misreadings. Actin, therefore, may not have been the most efficient housekeeping gene to use and if the qPCR experiment were repeated, two different housekeeping genes would have been used. For instance, *18S rRNA* (18S ribosomal RNA) and *TUBA* (*α*-tubulin), which are frequently used qPCR housekeeping genes (Kozera and Rapacz, 2013).

Real-time PCR was performed on all T2 transgenic lines to determine which lines were effectively overexpressing the respective transgene and to assess whether this increase affected the transcript level of the non-overexpressed DeSi protease. In the 60A, 60B and 60E transgenic lines, AT60 expression levels increased over a tenfold, whereas AT80 transcript levels decreased in 60B and 60E lines whilst increasing in the 60A line relative to WT. All AT80 overexpressing lines except for 80A, had double the AT80 expression level relative to WT, with all lines also increasing in AT60 transcript levels. The overexpressing lines with enhanced transgene expression levels were subjected to subsequent protein abundance assays. qPCR was also performed on the T3 AT60 and AT80 overexpressing lines that were propagated from the validated T2 gain-of-function mutants (60B, 60E, 80C and 80D), and also confirmed as homozygous transgenics through segregation analyses. Similar results were obtained, where the AT60 overexpressing lines displayed over a fourfold increase in AT60 mRNA abundance, along with increased AT80 levels. The two AT80 overexpressing lines also had significantly increased AT80 transcript levels as well as increased AT60 expression. This qPCR experiment confirmed the four T3 AT60 and AT80 transgenic lines as genuine overexpressing plants. This

finding would also mean that the observations from the phenotyping experiment conducted would not be solely attributable to the overexpression of one of the DeSi proteases and should be taken into consideration. However, as the expression level of the overexpressed SUMO protease were still greater than the non-overexpressed SUMO protease, the results could still be considered valid.

The real-time PCR results from both T2 and T3 overexpressing plants aligned with published literature on the transcriptional feedback observed between functionally redundant genes. The study by Chuang *et al.* (2012) established functional redundancy between two *pax*-like genes, *pax1* and *pax2*. The study found that the gene expression profiles of overexpressing Pax1 and Pax2 cells considerably overlapped with one another in the same direction. Whereas, the deletion or mutation of one gene resulted in minimal phenotypic difference due to the compensation by the other gene (Chuang *et al.*, 2012). In addition, the myogenic transcription factors in mice, *MyoD* and *Myf*-5, are also functionally redundant and display extensive cross-regulation. The deletion of *MyoD* consequently increased *Myf-5* transcription by threefold, whereas overexpressing *Myf-5* increased *MyoD* transcript levels (Pickett and Meeks-Wagner, 1995). Therefore, the qPCR results further substantiate the AT60 and AT80 DeSi proteases as functionally redundant proteins. The generation of single KO mutants and assessing functional compensation by the other DeSi protease resulting in minimal phenotypic differences to WT, would have further supported this notion.

Lastly, protein abundance assays were performed on the total protein extracted from the T2 overexpressing plants to confirm the presence and enhanced expression of the recombinant protein. The 60A, 60B and 60E transgenic lines had the highest level of the AT60 recombinant protein, aligning with the results from the qPCR experiment. All lines apart from 80A had enhanced expression of the AT80 recombinant protein. Despite 80B displaying strong protein expression levels, this line only had a slight increase in mRNA transcript levels and was not taken forward for propagation. The T2 AT60 and AT80 overexpressing lines that had the recombinant vector, displayed significantly enhanced expression levels of the respective transgene and were highly expressing the recombinant protein were selected for T3 propagation. The seeds from the individual plants of T2 60B, 60E, 80C and 80D transgenic lines, were screened on BASTA selection plates to identify the lines with 100% resistant seedlings, which were therefore considered homozygous overexpressing plants. The selected T3 overexpressing lines were confirmed via real-time PCR to have enhanced expression of the respective DeSi protease and were subsequently subjected to various phenotypic analyses.

4.5.2. Root length and fresh weight assays on AT60 and AT80 overexpressing lines further supports a role for AT60 and AT80 in the stress signalling pathways of Arabidopsis plants

The selected T3 lines, 60B-5, 60E-1, 80C-3 and 80D-6, were subjected to root length and fresh weight assays in normal and stress-induced conditions, replicating the experiment performed on the AT60-AT80KO mutants. The assay findings were compared to the phenotypes observed in the mutant plants to discern if the phenotypic changes were reversed. Under normal conditions, no significant differences were observed in the root length or biomass of the AT60-AT80KO lines. However, in stress-induced conditions, the AT60-AT80KO mutants were hypersensitive to the presence of both ABA and flg22 hindering their root elongation mechanism. This analysis would help further characterise the role the two DeSi proteases play in Arabidopsis development and stress responses.

In normal conditions, the root length assay found *ots1 and ots2* seedlings to have significantly shorter root lengths relative to all other lines except for the AT80 overexpressing line, 80C-3. Although 80C-3 had shorter root lengths than Col-0 and the other AT60 and AT80 overexpressing lines, this marginal difference was not significant. Consequently, this meant the single AT60 and AT80 overexpressing lines had the same primary root elongation rates as the WT. In contrast, the fresh weight assay revealed that all overexpressing lines were significantly lighter than Col-0. This indicated that the overexpressing lines may generally have slower or reduced overall growth in all plant tissues excluding root development of Arabidopsis plants. Therefore, the single AT60 and AT80 overexpressing lines had equivalent root elongation rates to WT, however, their overall growth were significantly slower in comparison to WT. Consequently, it could be assumed that the target substrates of AT60 and AT80 are involved in the growth and development of Arabidopsis.

In the root length assay, the presence of ABA significantly inhibited the root growth of all genotypes as expected and affected the AT60 and AT80 overexpressing lines the most. The root lengths were significantly shorter in 60E-1 and 80C-3 in comparison to Col-0. However, 60B-5 and 80D-6 lines had the same average root length as WT. These findings were reflected in the fresh weight assay which found a significant reduction in the biomass of ABA-exposed seedlings across all genotypes, with the greatest decrease in the AT80 overexpressing lines. Again, only the 60E-1 and 80C-3 overexpressing lines were significantly lighter than Col-0, whereas, the other two overexpressing lines had the same average biomass as WT in ABA-exposed seedlings. The variance observed in the ABA-induced growth inhibition across the AT60 and AT80 overexpressing lines could be explained by the variation of AT60 and AT80 transcript levels (Figure 4.3-B). The AT60 transcript level in 60B-5 was two times more than the 60E-1 line, whilst the AT80 expression level in 80D-6 was nearly double the levels detected in the 80C-3 line. Therefore, as 60B-5 and 80D-6 were significantly overexpressing their respective transgene more than their counterpart lines, these two lines could be considered the more dominant AT60 and AT80 overexpressing transgenic line. This would also explain the significant variation in biomass identified between 60B-5 and 80C-3 (Figure 4.6). As these two lines displayed the same degree of ABA-induced growth inhibition as Col-0, the phenotype in AT60 and AT80 overexpressing lines in response to ABA was the same as WT. Therefore, the overexpressing DeSi protease lines were not as susceptible to ABA-mediated root growth inhibition than WT.

The findings from analysing the AT60 and AT80 overexpressing lines aligned to a certain degree, with the phenotypes observed in the AT60-AT80KO mutants in response to ABA. As the KO lines were hypersensitive to ABA relative to WT, it would be expected the AT60 and AT80 overexpressing lines would exhibit increased tolerance to ABA in comparison to WT, attenuating ABA-mediated growth inhibition. This phenotype was observed in the Arabidopsis SUMO protease, OTS1. The study by Conti *et al.* (2008) displayed increased salt tolerance in OTS1 overexpressing lines in comparison to WT and the active site mutant ots1(C526S) protein. In the presence of salt, the OTS1 overexpressing lines considerably increased biomass production in addition to having significantly longer roots relative to Col-0 and the ots1(C526S) mutant. Although not entirely consistent with the 35S:OTS1 phenotype, the AT60 and AT80 overexpressing lines displayed no difference to WT in ABA-mediated root growth inhibition. This could be explained by the fact that joint overexpression of both DeSi proteases via the 35S promoter was necessary to display this phenotype (double overexpressing transgenics). Alternatively, co-suppression of the transgene could have occurred, where the increase in gene copy numbers reduces the degree of the transgene expression through endogenous silencing and only occurs

after the genes are actively transcribed (Smyth, 1997; Rajeevkumar *et al.*, 2015). Therefore, although mRNA transcript levels for the respective transgenes were high in the T3 overexpressing lines, the recombinant protein levels may not have correlated to this significantly enhanced expression. It would have been beneficial to check the recombinant protein expression levels of the T3 AT60 and AT80 overexpressing lines to confirm increased abundance of the respective DeSi proteases, as this may have provided an explanation for the inconsistency in results. The findings from these phenotypic analyses have provided further evidence to support the notion that both AT60 and AT80 could be involved in negatively regulating ABA signalling in Arabidopsis.

The presence of flg22 only significantly inhibited the root elongation of the two AT80 overexpressing lines. No differences in root growth between flg22-exposed and normally grown seedlings were detected in the AT60 overexpressing lines or Col-0. As the root length of the two AT80 overexpressing lines were significantly shorter than all other genotypes, this flg22-mediated root growth inhibition phenotype was unique to the AT80 overexpressing lines. The findings were consistent with the fresh weight assay which found that in the presence of flg22, the biomass of the AT80 overexpressing lines significantly decreased. Interestingly, both AT60 overexpressing lines increased in biomass when exposed to flg22, although the difference for only one line was significant. In comparison to WT, although AT60 overexpressing lines were heavier, albeit not at a significant amount, this increased biomass trend in response to flg22, was only observed in these transgenic lines. This suggests, the AT60 overexpressing lines are more tolerant to flg22 relative to WT, with no impact on the root elongation mechanism and even causing a slight increase in the growth of other plant tissues. In contrast, the biomass of the AT80 overexpressing seedlings considerably decreased in comparison to all other genotypes. Consistent with the root length assay, this implies AT80 gain-of-function mutants are considerably more susceptible to flg22-mediated growth inhibition relative to WT, resulting in the overall global growth arrest of all plant tissues as well as root development.

The results from phenotyping the AT60 overexpressing lines in response to flg22, align with the phenotypes observed in the AT60-AT80KO mutants. The KO mutant lines were hypersensitive to flg22, exhibiting considerable inhibition in root growth relative to WT. The expected opposite phenotype was observed in the AT60 overexpressing lines, which showed slight tolerance to the pathogen response elicitor. The AT60 gain-of-function mutants in response to flg22 had marginally longer root lengths and heavier biomass than Col-0 and the overexpressing lines showed no difference in the root lengths of seedlings grown in normal and flg22-induced conditions. These results were consistent with the Conti *et al.* (2008) study, where OTS1 overexpressing lines had greater biomass production and significantly longer roots than WT. Furthermore, this phenotype was also observed in findings from a thesis study that investigated the Arabidopsis DeSi protease, Desi3a (Yates, 2018). Both this study and the paper by Orosa *et al.* (2018) demonstrated the presence of flg22 caused significant root growth inhibition in Col-0 and even more restricted root elongation in *desi3a-1*, coinciding with the phenotype observed in the AT60-AT80KO mutant. In contrast, the Desi3a overexpressing lines appeared to be insensitive to flg22-treatment, with flg22-exposed seedlings growing at the same rate, in terms of root length and biomass, as seedlings grown in normal conditions (Yates, 2018). This aligned with the phenotype observed in the AT60 overexpressing line, therefore confirming the AT60 protease plays a significant role in pathogen response and most likely negatively regulates plant immune responses.

In contrast, the phenotype of the AT80 overexpressing transgenic line in response to flg22 did not align with the AT60-AT80KO mutant phenotype or published literature on Desi3a. The AT80 gain-of-function mutants were considerably more susceptible to flg22 relative to WT, consequently resulting in decreased biomass production and root growth restriction. Co-suppression of the transgene could again explain this phenomenon. Transgenic plants overexpressing a particular gene have been found to display individual variation in the expression or silencing of the transgene (Matzke *et al.*, 2009). With the presence of the native AT80 gene in the background and the overexpressing AT80 transgene in the transgenic lines, endogenous silencing could have occurred. Despite the very high AT80 transcript levels identified in the gain-of-function mutants, as co-suppression only takes place post-transcription, the recombinant protein levels may not have corresponded to this elevated expression (Smyth, 1997; Rajeevkumar *et al*., 2015). Endogenous silencing would explain why the AT80 overexpressing transgenic lines were displaying the same phenotypes observed in the AT60-AT80KO mutant lines in response to flg22. Therefore, it would have been beneficial to have analysed the protein levels of the AT80 overexpressing lines as well as to generate a complementation transgenic line by overexpressing AT80 in the KO mutant background. These findings were not able to further support the theory that the AT80 protease is involved in flagellin sensing and could be negatively regulating plant immune responses.

4.5.3. Localisation studies suggest the overexpressed AT60 and AT80 proteases localise to the cell periphery and their spatial expression pattern and level significantly changed in response to ABA and flg22

Confocal microscopy studies on the T3 AT60 and AT80 overexpressing lines were used to provide better insight to the organ-specific and subcellular localisation of the two recombinant DeSi proteases in Arabidopsis. The stable transformation and expression of the overexpressing vector in Arabidopsis plants would provide more accurate findings than the results from the transient expression assay conducted in *N. benthamiana* plants. However, as the expression of the proteases are driven by a non-native promoter, the experiment findings should not be considered reflective of the endogenous sub-cellular localisation of AT60 and AT80. The localisation study of the YFP-tagged DeSi proteases was first performed in normal conditions to identify the subcellular localisation of the overexpressed AT60 and AT80 protein across the Arabidopsis seedling. The localisation studies were then repeated in the presence of ABA and flg22, to determine the spatial expression patterns of AT60 and AT80 in response to the stress inducers. The findings from this study could be compared with the qPCR experiment results and the stress-induced phenotypic assays to further elucidate the function of AT60 and AT80 in Arabidopsis development and defence responses.

The overall expression of AT60 in AT60 overexpressing lines across the four plant organs was generally lower than AT80 protein expression in AT80 overexpressing lines. AT60 was absent from the nucleus and could be faintly localised to the cell periphery, most likely the plasma membrane, in the epidermal leaf and root cells. The AT80 protease was highly expressed in the root and root tip cells and was also localised outside the nucleus towards the plasma membrane, although there was difficulty in compartmentalising the protein expression to a particular subcellular component. Although AT80 appeared to localise in the guard cell, this was probably the autofluorescence emitting from chlorophyll in the guard cell. This was most likely the case as AT80 had

very low fluorescence in the leaf epidermal cells and chlorophyll fluorescence can be detected at wavelengths close to the 514nm laser used for this study (600-660nm) (Schulte *et al.*, 2006).

Interestingly, these results align with published literature on the localisation of DeSi proteases. The DeSi2 member proteins in humans, which are orthologous to AT60 and AT80 proteases (Yates, 2018), have been found to localise mainly in the cytoplasm along the periphery of the cell (Shin *et al.*, 2012). Both studies by Orosa *et al.* (2018) and Yates (2018), have also conclusively determined the Desi3a protease localises to the membrane fraction of the cell with no evidence of nuclear localisation. Furthermore, Yates (2018) conducted preliminary localisation studies on the AT60 and AT80 DeSi proteases confirming no nuclear localisation was observed in the AT80 protease, whereas AT60 was faintly expressed along the plasma membrane. In contrast, ULP SUMO protease family members in Arabidopsis were found to primarily localise in the nucleus including the SPF1/2 and OTS1/2 proteases (Conti *et al.*, 2009; Liu *et al.*, 2017). Therefore, it can be postulated through this localisation study, qPCR experiment and published literature, that both AT60 and AT80 DeSi proteases appear to localise outside the nucleus in the cell periphery, most likely along the plasma membrane. The AT80 protease is expressed significantly more in the roots relative to AT60, whilst the AT60 protease is relatively more expressed in the leaves. However, to conclusively determine the endogenous sub-cellular localisation of the two DeSi proteases, the transgenes would need to be expressed under their respective native promoter instead of the constitutive, 35S promoter.

In ABA-induced conditions, the overall expression of AT60 significantly decreased relative to normal AT60 levels, especially in the root system were hardly any fluorescence signals were detected. Although AT60 expression significantly reduced in the leaf cells, the spatial expression pattern did not change in the plant stem tissue, with AT60 faintly localising to the plasma membrane. These findings were consistent with the qPCR experiment which found AT60 expression to significantly reduce in the presence of ABA. It could therefore be postulated that in ABA-induced conditions, the level of AT60 rapidly decreases in Arabidopsis WT plants, similar to the OTS1 protease levels in response to salt stress (Conti *et al.*, 2008). However, as the AT60 protease was linked to a strong constitutive promoter the observed difference in expression level could instead be attributed to changes in localisation or protein degradation, rather than the presence of ABA. Repeating this experiment with the expression of the AT60 protease driven under its endogenous promoter would be able to clarify this uncertainty.

In contrast, AT80 protein levels significantly increased across all plant organs in response to ABA. This was consistent with the findings from the real-time PCR. The subcellular localisation of AT80 remained consistent to the spatial expression pattern observed in normal conditions, with the DeSi protease localising to the plasma membrane across all four plant tissues. AT80 expression levels significantly increased the most in the root tip of the Arabidopsis seedling in the presence of ABA, relative to other plant organs. These findings suggest the target substrate of AT80 most likely localises in the plasma membrane of root tip cells in Arabidopsis plants. It could, therefore, be proposed that in response to ABA, AT80 expression increases particularly in the root tip, consequently reducing the abundance of the SUMO-conjugated AT80-target substrate related to root elongation, thereby activating ABA-mediated root growth inhibition. On that account, these findings showcase functional divergence between the two highly homologous DeSi proteases in response to ABA, where AT80 increases in expression, whilst AT60 decreases. Functional divergence in response to stress has been observed

between pairs of gene duplicates and homologous genes in Arabidopsis (Zou *et al.*, 2009). For instance, one study demonstrated the partially redundant function between the *OsbHLH068* and *AtbHLH112* gene in mediating abiotic stress in Arabidopsis, whilst for the control in flowering, they functioned antagonistically, presumably due to divergent evolution (Chen *et al.*, 2017). In this case specifically, it could be theorised there is unequal functional redundancy between the two DeSi proteases as seen in SPF1/2 (Castro *et al.*, 2018), tending towards AT80 as being more important. AT80 is expressed significantly more across plant organs relative to AT60 and has a more prominent response to stress stimuli, mirroring the characteristics observed in SPF1; the more dominant of the SPF1/2 proteases.

There are two theories that could link the AT80 DeSi protease with ABA signalling. The first is that AT80 positively regulates ABA signalling. In response to ABA, the discerned upregulation of the AT80 protease coincides with findings of the SPF1 protease, which has demonstrated its SUMO protease activity in positively regulating ABA signalling during early seedling development. This recent study established that ABA promotes SPF1 protein abundance and postulated SPF1 subsequently adjusts the protein abundance of two transcription factors regulating ABA, which are SUMOylated by SIZ1; ABI5 and MYB30 (Wang *et al.*, 2018). SIZ1 represses ABA signalling through mediating the SUMOylation of ABI5, a positive ABA regulator, reducing ABI5 activity, and SUMOylating MYB30, a negative regulator of ABA, promoting MYB30 stability and activity (Miura *et al.*, 2009; Zheng *et al*., 2012). ABA-induced SPF1 accumulation could increase ABI5 deSUMOylation, activating ABI5 to induce ABA-mediated growth arrest. Whereas, enhanced MYB30 deSUMOylation, allows the ubiquitin E3 ligase, RHA2b, to target MYB30 for degradation (Zheng *et al.*, 2018), deactivating the negative regulator of ABA signalling. This theory was supported by the observation that *spf1* mutants exhibited an ABA-insensitive phenotype in root elongation (Wang *et al.*, 2018). Although these findings contradict the observed phenotype in the AT60-AT80KO mutant root length assay, this variation could be attributed to the AT60 DeSi protease KO. Single KO mutants could have elucidated this suspicion. Nonetheless, it could, therefore, be postulated that the AT80 DeSi protease could be functioning similarly to the SPF1 protease in positively regulating ABA signalling, and even targeting the same or similar functioning transcription factors or proteins, most likely localised in the Arabidopsis root system.

Alternatively, AT80 could be negatively regulating ABA signalling as observed in the AT60 and OTS1 protease (Conti *et al.*, 2008). Contradictory to the response AT60 and OTS1 have to ABA, the expression of AT80 significantly increases in response to ABA. This would consequently enhance deSUMOylation of the AT80-target protein, reducing the abundance of the SUMOylated substrate leading to the ABA-mediated growth repression phenotype. According to the phenotypes observed in this study, AT80 overexpressing lines displayed no difference in ABA-induced growth inhibition to WT, whereas AT60-AT80KO mutants were hypersensitive to ABA, significantly enhancing the ABA-induced growth repression phenotype. If AT80 was negatively regulating ABA, this would indicate the concentration of the SUMOylated AT80-target protein could be regulating the ABA-mediated growth inhibition phenotype. For instance, in AT80KO, SUMOylated levels of the AT80-target substrate would be very high resulting in strong ABA-induced growth arrest, whereas AT80 overexpression results in very low SUMOylated levels of the AT80-target substrate leading to a suppressed ABA-mediated growth inhibition phenotype. Therefore, further studies on AT80 transgenic lines are required to elucidate the role AT80 plays in ABA signalling.

On the other hand, in response to flg22, the expression of the AT60 protease significantly reduced in the root cells, root tips and leaves of Arabidopsis plants. Whereas, no difference in AT60 expression was detected in the stem. Even under flg22 exposure, the AT60 protein seemed to localise to the outside of the cell, most probably the plasma membrane. in all analysed plant tissues. Interestingly, in the leaf epidermis, AT60 expression could also be observed in the cytoplasm which was not detected in normal conditions. This subcellular localisation change could be attributable to the presence of flg22 inducing a localisation response or simply due to the transgene's strong constitutive promoter. These findings were consistent with the qPCR results as well as the published literature results on Desi3a. Yates (2018) found the transcript and protein levels of Desi3a significantly decreased in Col-0 plants following flg22-treatment. This observation was also seen in the Orosa *et al.* (2018) paper where rapid degradation of Desi3a occurred within 10 minutes of flg22 treatment. Both studies established Desi3a as a negative regulator in immune signalling, where flg22-induced rapid degradation of Desi3a promotes the accumulation of SUMOylated FLS2, a known target of Desi3a, subsequently activating immune signalling in plants. With the results from this study aligning with the Desi3a-1 published findings, it could be concluded that the AT60 DeSi protease acts as a negative regulator in plant immune responses. Upon pathogen perception, AT60 protein levels significantly decrease, leading to the accumulation of SUMOylated AT60-target substrates, most likely situated in the plasma membrane, consequently activating early cellular immune responses to the pathogen attack.

The expression of the AT80 protease in response to flg22 considerably reduced or remained constant to normal conditions in the root system, whereas expression significantly increased in the leaf and stem tissue. Interestingly, in the presence of flg22, AT80 localised to the plasma membrane across all plant organs, however in the leaf epidermis, AT80 protein accumulation could be detected in the guard cells. In this case, the distinctive fluorescence spatial expression pattern displayed in the guard cell (Figure 4.9) appears to represent the recombinant AT80 protease and not chlorophyll autofluorescence. These findings were to an extent, consistent with the qPCR findings and the speculated model on AT60 and Desi3a, where upon flg22 perception, AT80 protein abundance decreases to activate immune responses, thereby, negatively regulating pathogen responses. However, similar to the response to ABA, AT80 functionally diverges from AT60 during pathogen infections as AT80 accumulates in the guard cell upon flg22 perception. This indicates AT80 may have a role in regulating stomatal aperture in response to pathogen attacks. Supposedly, the AT80 DeSi protease accumulates in the guard cell upon flg22 perception, to deSUMOylate a defence response regulator of stomatal aperture. In response to bacteria and PAMPs, the stomata in Arabidopsis plants have been proven to close as part of PTI. Zhang *et al.* (2018) demonstrated how the PAMP, flg22, triggers stomatal closure inhibiting light-triggered stomatal opening in an FLS2-dependent manner. FLS2 has been localised in the guard cell (Beck *et al.*, 2014), is a proven target substrate for Desi3a-mediated deSUMOylation and FLS2 hyper-SUMOylation triggers intracellular immune signalling (Orosa *et al.,* 2018). Therefore, it could be postulated that AT80 increases in abundance within the guard cell upon flg22 perception to target FLS2 or another defence response elicitor to induce stomatal closure as part of PTI. Future studies are necessary to further explore the potential role AT80 may have in PTI-associated guard cell responses.

4.5.4. Future work in relation to this study

There are several factors and experiments which should be conducted to further elucidate the function and characteristics of the AT60 and AT80 DeSi proteases. Performing a protein abundance to check the recombinant protein levels of the T3 overexpressing transgenic lines would have been imperative to understanding the findings, as well as in explaining the inconsistency in results obtained from this experiment. More phenotyping experiments on the AT60 and AT80 overexpressing lines would have provided more insights into the function of the two DeSi proteases in Arabidopsis development and immune signalling. This could include conducting germination rate assays, flowering time experiment and a Pst infection assay to determine if either AT60 or AT80 had longer-term infection impacts. For the confocal microscopy experiments, the implementation of necessary cellular markers would have validated the compartmentalisation of the DeSi proteases. For instance, fluorescently labelled lectins like wheat germ agglutinin or CellLight reagents could be used to label the plasma membrane and ReadyProbes reagents for the fluorescent staining of nucleic acids. In addition, optimisation of the confocal microscope settings to spectrally sperate total YFP fluorescence from chlorophyll autofluorescence (Watkins *et al*., 2014) would have improved the accuracy in determining the subcellular localisation of the AT60 and AT80 proteins. It would have also been beneficial to have cloned the promoter region of both DeSi proteases, so the transgenes were expressed under their native promoter instead of the constitutive 35S promoter. This would have displayed the endogenous protein expression levels and subcellular localisations of the two DeSi proteases in Arabidopsis. The generation of AT60 and AT80 double overexpressing lines would have determined whether the phenotypes observed in the single overexpressing plants were attributable to the gain-of-function in one of the DeSi proteases and not both. More importantly, the creation and analysis of single KO mutants would have elucidated the many suspicions and theories in question that had arisen from this study regarding the characteristics and functionality of AT60 and AT80. In-depth analyses on the single KO mutants would have confirmed the functionally redundant or divergent nature of the two highly homologous DeSi proteases and validated the regulatory role each protease has in response to ABA and flg22. Lastly, the generation of a complementation line by overexpressing the DeSi protease in the KO mutant background would have verified whether the DeSi proteases were solely responsible for the phenotypes observed in the mutant lines confirming their functionality.

4.5.5. Final concluding remarks

This chapter first justifies how the T3 AT60 and AT80 overexpressing lines were generated and validated as gain-of-function mutants. The overexpressing lines which carried the recombinant vector, exhibited enhanced transcript levels of the respective protease, displayed heightened protein levels and screened as homozygous transgenics, were subjected to the same phenotyping experiments performed on the KO mutants. In normal growth conditions, AT60 and AT80 overexpressing lines exhibited the same root elongation rate as WT although slower overall growth was detected. In response to ABA, both AT60 and AT80 were not as susceptible to ABA-mediated growth inhibition than WT. Interestingly, in the presence of flg22, AT60 overexpressing lines were slightly more tolerant to the pathogen response elicitor, whereas AT80 overexpressing lines, were hypersensitive to flg22 most likely as a result of transgene co-suppression. Confocal microscopy studies on the overexpressing lines found both AT60 and AT80 localising outside the nucleus towards the cell periphery, presumably in the plasma membrane, across the leaf, stem and root tissue of the Arabidopsis seedling. This implies the target substrate of both DeSi proteases are most likely localised in the

plasma membrane or cell periphery. In both normal and stress-induced conditions, AT80 had significantly higher overall protein expression especially in the root system, suggesting unequal functional redundancy may exist between the two highly homologous proteases, tending towards AT80 being more important. In the presence of both ABA and flg22, AT60 expression significantly decreased universally, aligning with published literature findings on OTS1/2. Therefore, looking at all findings obtained from this study, it could be postulated that AT60 plays a significant role in ABA signalling and plant immune responses, most likely as a negative regulator. On the other hand, AT80 protein expression considerably increased in response to ABA, substantiating the theory of unequal redundancy existing between AT60 and AT80. The findings suggest AT80 is implicated in ABA signalling, however, whether AT80 was positively or negatively regulating the ABA pathway could not be hypothesised. In response to flg22, AT80 protein expression reduced in the roots, whilst increasing in the leaves, specifically the guard cell. These findings propose a role for AT80 in the immune signalling response and postulated AT80 could even be implicated in PTI-associated guard cell responses.

5. Investigation of the Biochemical Properties of the AT60 DeSi Protease

5.1. Introduction

DeSUMOylating proteases are crucial in maintaining the equilibrium in the SUMO signalling pathway, providing reversibility and flexibility. These specialised proteases cleave the isopeptide bond exactly between the glycine residue of the SUMO protein and the cognate substrate, freeing the SUMO moiety for subsequent conjugation cycles. Previous reports have demonstrated the SUMO protease activity *in vitro* for a limited number of the bona fide SUMO proteases including ESD4 and OTS1/OTS2 (Colby *et al.*, 2006; Conti *et al.*, 2008, 2014). Furthermore, these studies confirmed that the SUMO protease activity requires the specific cysteine active site, as mutating this residue to a serine results in the abolition of the protease activity. Within the newly-identified class of Arabidopsis SUMO proteases, DeSi proteases, which act only in the removal of SUMO conjugates, only one Desi SUMO protease, Desi3a, has been described and biochemically characterised to date. The study by Orosa *et al.* (2018), conducted an *in vitro* deSUMOylation assay that confirmed Desi3a as a bona fide DeSi SUMO protease with the ability to remove isopeptide-linked SUMO proteins from the target substrate, FLS2.

The aim of this chapter was to biochemically characterise one of the Arabidopsis DeSi proteases, AT60, and demonstrate its SUMO protease activity. To further investigate the catalytic function of the DeSi motif, a mutant AT60 protein was generated with the catalytic core cysteine residue substituted for the structurally similar but chemically different amino acid, serine. In order to achieve this, the optimal expression conditions for the recombinant AT60 protein in *E. coli* cells were first established. Subsequently, the AT60 WT and mutant protein, as well as the target substrate to be tested, the JASMONATE ZIM-domain (JAZ) 6 protein, were all optimally expressed and purified. To ascertain whether the putative DeSi protease could remove SUMO from the JAZ6 target substrate, an *in vitro* deSUMOylation assay was performed. Unfortunately, due to time constraints, this experiment was not fully completed, and the assay was only repeated twice. If more time was available, the experiment would have been repeated several times for the protocol to be optimised and the SUMO protease activity of the AT80 DeSi protease would have also been investigated.

5.2. Expression of Recombinant AT60, AT60^{C115S} and JAZ6 Proteins in *E. coli*

DH5α E. coli cells overexpressing the AT60 gene within the entry vector pENTR/D-TOPO, was kindly donated by Dr. Orosa. The LR reaction using the Gateway™ LR Clonase™ Enzyme mix from Invitrogen was then performed to transfer the cloned AT60 cDNA fragment into the destination vector, pDEST17. This destination vector consists of an N-terminal His tag and a selection cassette for ampicillin (carbenicillin), allowing positive selection on antibiotic-containing media. Furthermore, the pDEST17 vector is a protein expression system that is activated in the presence of IPTG, which can increase the cellular concentration of the protein of interest by x1,000 times (Lewis and Bell, 2000). The AT60 gene within the pDEST17 vector was transformed from DH5α *E. coli* cells into competent BL21 (DE3)-RIL cells. This *E. coli* strain contains a vector with tDNA codons from eukaryotic systems optimised for enhancing the expression of eukaryotic proteins in bacterial organisms (Novy *et al*., 2001). For the *in vitro* deSUMOylation assay, *E. coli* SS+ cells expressing the JAZ6 protein in

the pDEST17 vector were kindly donated by Dr. Srivastava. These *E.coli* SS+ cells comprise of Arabidopsis SUMO conjugation machinery proteins with SUMO1 modified to expose the C-terminal GG sequence (Okada *et al.*, 2009). Therefore, they would ensure the JAZ6 protein, which has been proven to be SUMOylated in Arabidopsis (Srivastava *et al.*, 2018), would be heavily SUMOylated, therefore acting as an optimal SUMOylated target substrate for the deSUMOylation assay. In order to prove the active site cysteine residue and cease protease activity, a mutant AT60 gene was generated using site-directed mutagenesis.

5.2.1. Generation of the AT60^{C115S} Mutant Protein

Plasmid DNA isolated from the DH5α *E. coli* cells overexpressing AT60 in the entry vector pENTR/D-TOPO was used to produce the mutant AT60 protein. Site-directed mutagenesis was performed to mutate the active site cysteine residue at the amino acid position 115 of the AT60 gene (Figure 5.1), to the structurally similar but chemically different amino acid serine. This was accomplished by changing the guanine nucleotide base to a cytosine, altering the DNA sequence from TGC to TCC, therefore resulting in the substitution of cysteine to serine. Sequencing was performed on the mutated plasmid to confirm the site-directed mutagenesis was successful. Subsequently, the LR clonase reaction was conducted to transfer the cloned AT60^{C115S} mutant cDNA fragment into the destination vector, pDEST17. The mutated AT60^{C115S} gene within the pDEST17 vector was then transformed into BL21 (DE3)-RIL cells for protein expression analysis.

AT60	113	K	N	C	N	H	117
		AAA	AAT	TGC	AAT	CAC	
AT60^{C115S}	113	K	N	S	N	H	117
		AAA	AAT	TCC	AAT	CAC	

Figure 5.1 – Coding and protein sequence of the active site of AT60 WT and AT60^{C115S} mutant. The active site cysteine (amino acid position 115) and mutated serine are highlighted in red.

5.2.2. Protein Expression of the AT60 WT, AT60^{C115S} Mutant and JAZ6 Protein

With the WT and mutated AT60 gene fragments in the pDEST17 vector and BL21 (DE3)-RIL cells, the recombinant *E. coli* cells, including the SS+ cells expressing the JAZ6 protein (pDEST17), were ready to be tested for expression optimisation. Each recombinant *E. coli* strain carrying the AT60 WT, AT60^{C115S} mutant and JAZ6 gene, were expressed using the protein expression protocol. Numerous experiment attempts found the AT60 protein absent from the soluble fraction when expressed at 28°C. Therefore, AT60 was expressed overnight at a cooler temperature of 18°C and half the standard concentration of IPTG (0.5mM) as an attempt to slow down its translation and improve protein folding. The conditions under which the recombinant proteins were best expressed were recorded and can be observed in Table 5.1. The protein expression analysis results are displayed in Figure 5.2.

Table 5.1 – Optimal expression conditions for the recombinant proteins

Cell Line	Recombinant Protein	Optimal Expression Temperature	Optimal Induction Time	IPTG Concentration (mM)
E. coli SS+ cells	JAZ6	30°C	3 hours	1
BL21 (DE3)-RIL	AT60	18°C	16 hours (overnight)	0.5

BL21 (DE3)-RIL	AT60^{C115S}	18°C	16 hours (overnight)	0.5

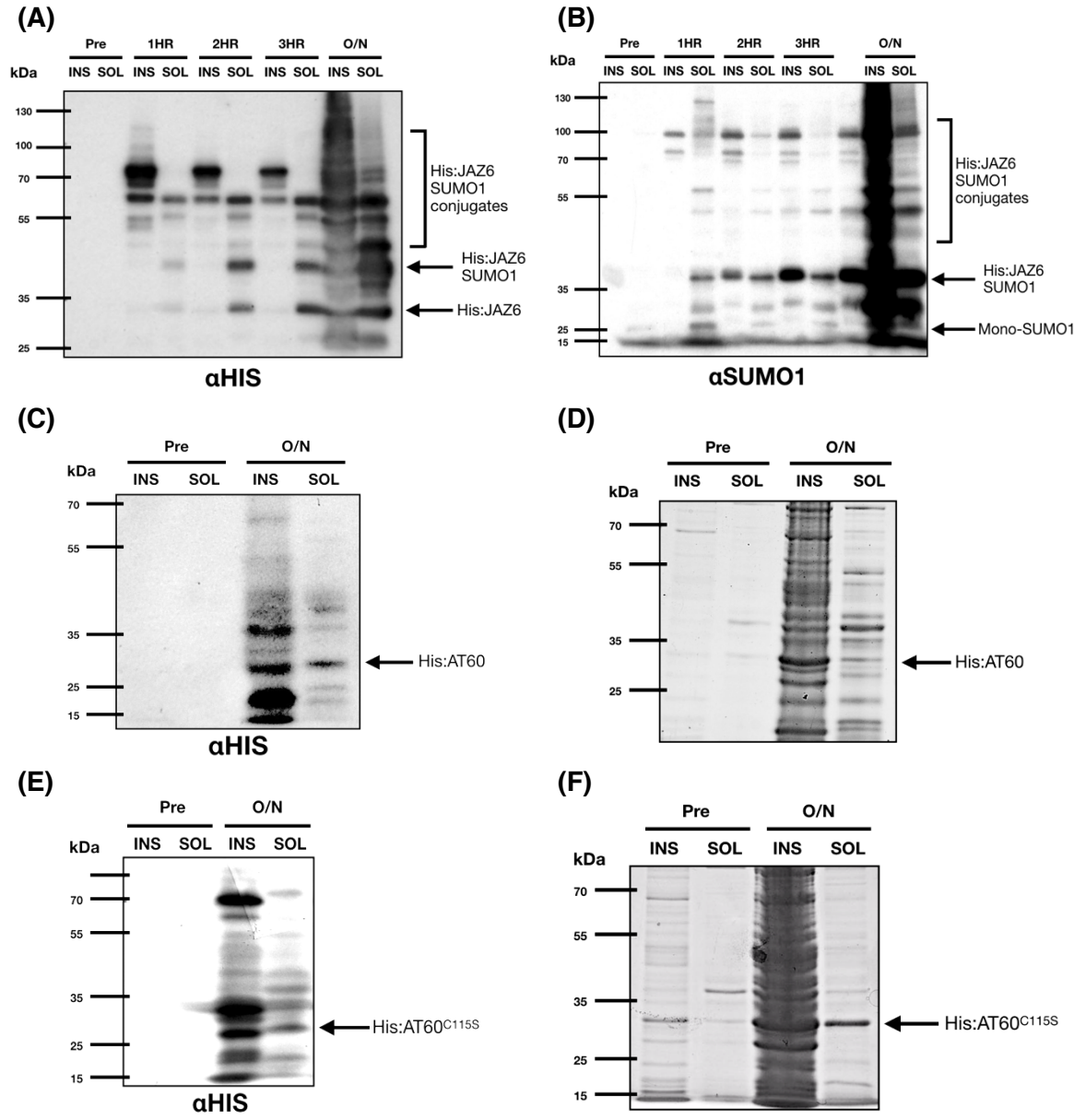


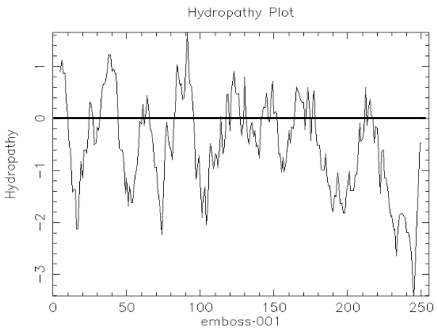
Figure 5.2 – AT60, AT60^{C115S} and JAZ6 protein expression. Transgenic *E. coli* strains were induced with 0.5-1mM IPTG and incubated at their respective optimal expression temperature and induction time. Post-induction samples were taken every hour for 3 hours for the JAZ6 protein (A-B) or after 16 hours overnight (O/N) for the AT60 protein (C-F). Samples from pre-induction (Pre) and post-induction were then processed into soluble (SOL) and insoluble (INS) fractions before being separated by SDS-PAGE. The SDS-PAGE gels were visualised through a western blot (A, B, C and E) probed with anti-His (A, C-F) and anti-AtSUMO1 antibodies (B), and a Coomassie stain (D and F). A) Western blot probed with anti-His to detect His-fusion proteins, displays the His:JAZ6 protein band at ~31kDa and SUMOylated His:JAZ6 at ~43kDa in the soluble fraction. The heavier protein bands forming a ladder pattern most likely represent poly-SUMO1 chains conjugated to the His:JAZ6 protein. B) Western blot probed with anti-AtSUMO1 antibodies to detect SUMO chains, reveals the presence of a SUMO1 monomer at ~12kDa in the soluble fraction and SUMOylated

His:JAZ6 at the expected molecular mass of ~43kDa in the soluble and insoluble fraction, whilst the higher bands likely represent SUMO1 multimers and poly-SUMO1 chains conjugated to His:JAZ6. C and E) Western blot shows AT60 (C) and AT60^{C115S} (E) at ~30kDa in the soluble fraction. D and F) Coomassie stain displays AT60 (D) and AT60^{C115S} (F) at ~30kDa in the soluble fraction. His:JAZ6 = 30.85kDa and His:AT60/AT60^{C115S} = 29.12kDa.

As expected, the accumulation of the respective recombinant proteins increased after IPTG induction across all three proteins. Post-induction, the expression of the JAZ6 protein heightened over time and was found to be more prominently present in the soluble fractions (Figure 5.2-A and Figure 5.2-B). The heavier protein bands appearing as a ladder in the soluble fraction indicate JAZ6 has been SUMOylated. As SUMO1 proteins are approximately 11.5 kDa, mono-SUMOylation of JAZ6 would result in a band at approximately ~43kDa, which correlates to the next band above the His:JAZ6 protein band. The following heavier bands forming a ladder pattern observed in Figure 5.2-A and Figure 5.2-B, reflect the possible conjugation of multiple SUMO proteins to the JAZ6 substrate resulting in a poly-SUMO1 chain. This assay confirms the expression and SUMOylation of JAZ6, as well as the generation of a SUMO conjugation ladder ideal for the *in vitro* deSUMOylation assay. The AT60 and AT60^{C115S} protein proved difficult to express and was prominently in the insoluble fraction concentrated in inclusion bodies. However, growing the recombinant bacterial strain at 18°C for 16 hours (overnight) to improve protein folding, resulted in a small proportion of the proteins accumulating in the soluble fraction, as well as the insoluble fraction, as can be seen in the western blot (Figure 5.2-C and Figure 5.2-E) and Coomassie stain (Figure 5.2-D and Figure 5.2-F). Both His-tagged AT60 and AT60^{C115S} proteins can be found at the expected molecular mass of approximately ~30kDa. The lower protein bands most likely represent breakdown products of the His-AT60 and His-AT60^{C115S} protein highlighting the instability of the DeSi protease. The other protein bands observed in the blot are most probably a product of non-specific anti-His binding. Despite only a small proportion of the AT60 protein being present in the soluble fraction, the optimal expression conditions for the AT60, AT60^{C115S} and JAZ6 recombinant bacterial strains were used for purification.

In attempt to explain the abundance of AT60 in the insoluble fraction, a hydropathy plot was constructed for the AT60 DeSi protease. As displayed in Figure 5.3, seven peaks were clearly hydrophobic accounting for approximately ~40% of the total peptide sequence. As nearly half of the AT60 protease scores positively on the hydropathy index and has hydrophobic properties, this provides an explanation for its abundance in the insoluble fraction. The isoelectric point was also considered as a determining factor. With the isoelectric point of the AT60 protease at 8.12 and the pH of the *E. coli* cytoplasm estimated around 7.2-7.8 (Wilks and Slonczewski, 2007), the AT60 protease will have less charged residues and would be closer to net-neutrality, consequently limiting the protein's solubility, resulting in protein aggregation. Therefore, the hydrophobicity and isoelectric point of the AT60 protease provide an explanation for the minimal accumulation of the protein in the soluble fraction.

Figure 5.3 – Hydropathy plot of AT60 DeSi protease. To produce the hydropathy plot, the respective protein sequence was obtained from TAIR database and the sequence was input into the online EMBOSS Pepwindow software to develop the hydropathy plot (https://www.ebi.ac.uk/Tools/seqstats/emboss_pepwindow/)



5.3. Purification of *E. coli* cells expressing recombinant AT60, AT60^{C115S} and JAZ6 proteins

The recombinant proteins were expressed optimally and prepared for purification by lysing and centrifuging the bacterial cells for extraction of the supernatant as the soluble fraction. This supernatant was used as the starting solution for the protein purification experiment. As all three proteins were tagged with an N-terminal His tag (pDEST17), the HisTrap column was used for the purification of the proteins. The cell extracts were run through the HisTrap column attached to an AKTA machine, enrichening it with the respective proteins and subsequently eluting it out into fractions comprising of varying protein content. UV absorbance levels of the sample flowing through the column measured by the AKTA machine indicated which elution fraction contains the highest concentration of the respective protein. The collected fractions with the highest protein concentration were then analysed on SDS-PAGE and visualised with a western blot to confirm the presence of the recombinant protein. Figure 5.4 displays the results from the purification analysis.

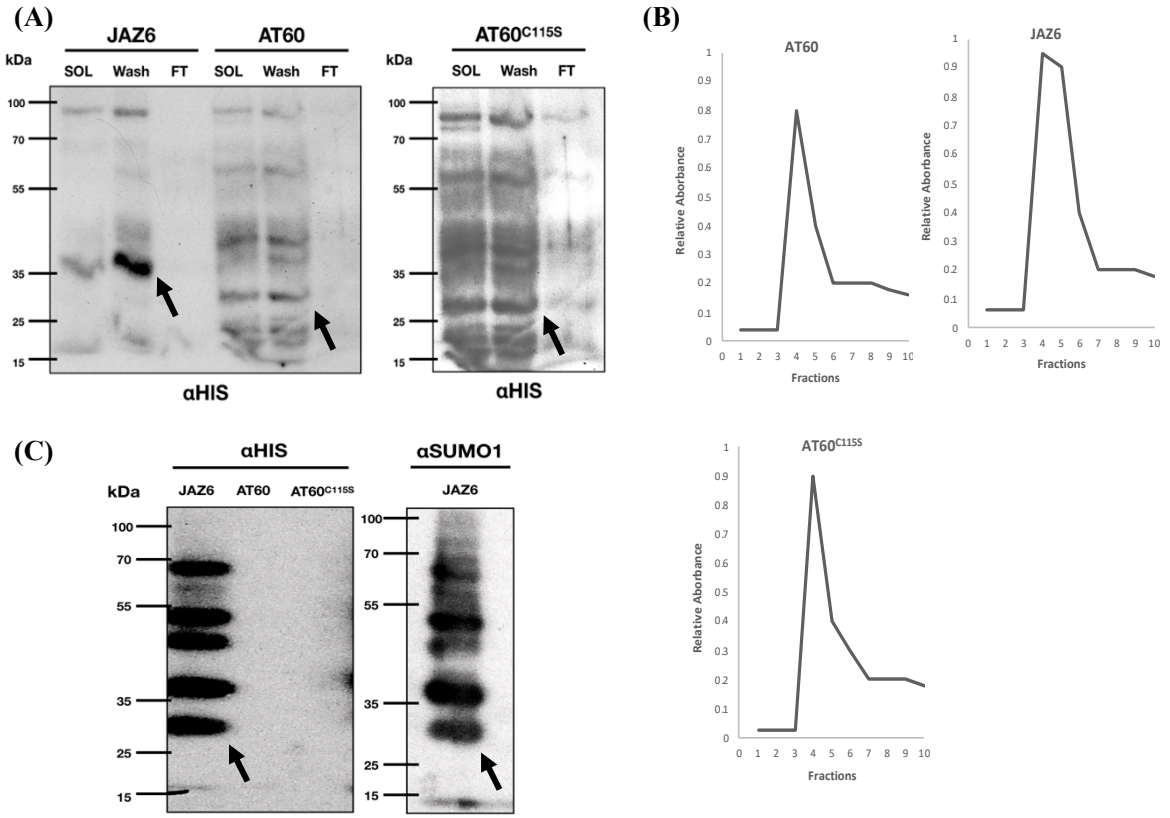


Figure 5.4 – Purification analysis of recombinant JAZ6, AT60 and AT60^{C115S} proteins. The recombinant bacterial strains were grown at their respective optimal expression conditions for cell harvesting. The soluble cell extract was then extracted and passed through a HisTrap column for purification. The pre-purification

soluble lysate (SOL), wash fractions (wash) and lysate flow through (FT) fractions were obtained as controls prior to the elution. The arrows denote the corresponding recombinant proteins at their respective molecular weight. A) Control samples comprising of the soluble lysate, wash fraction and lysate flow through for each recombinant protein were separated by SDS-PAGE and visualised with a western blot probed with anti-His. B) UV absorbance levels measured for each recombinant protein sample, revealed which elution fractions contained the highest protein concentration (JAZ6 = 3-5; AT60 = 4-6; AT60^{C115S} = 4-6). C) Based on the UV absorbance level, a small proportion of the fraction containing the highest protein concentration for each respective protein (JAZ6, AT60 and AT60^{C115S} = 4) was run on SDS-PAGE and visualised with a western blot probed with anti-His for all three recombinant proteins and anti-AtSUMO1 for JAZ6. His:JAZ6 = 30.85kDa and His:AT60/AT60^{C115S} = 29.12kDa.

In Figure 5.4-A, the control fractions for each of the recombinant proteins obtained during the purification experiment were tested. The control samples comprised of the soluble cell extract, the flow through from the lysate and the flow through collected during the binding buffer washes. The arrows denote the presence of all three of the respective proteins in the binding buffer wash fraction, indicating the proteins were not captured efficiently in the HisTrap column. Figure 5.4-B displays the UV absorbance level measured for each of the recombinant protein samples during elution and reveals which elution fraction contained the highest concentration of the respective protein. In Figure 5.4-C, a small proportion of the highest protein content fraction for each recombinant protein was separated on SDS-PAGE. Immunoblot analysis with anti-His and anti-AtSUMO1 revealed the presence of His:JAZ6 in the elution fraction at ~30kDa and SUMOylated His:JAZ6 at ~43kDa, however, there was no presence of the AT60 and AT60^{C115S} protein. This would be explained by the findings from Figure 5.3-A where the proteins were found in the wash fractions as they had not been properly captured by the HisTrap column. Despite there being no detection of the WT AT60 and AT60^{C115S} protein, the fractions with the highest protein concentrations for each protein sample (JAZ6 = 3-5; AT60 = 4-6; AT60^{C115S} = 4-6) were pooled together for dialysis and concentration with the aim of increasing the protein concentrations for the *in vitro* deSUMOylation assay.

5.4. In *vitro* DeSUMOylation Assay to Test the AT60 SUMO Protease Activity on JAZ6

Once all components had been purified, the proteins were dialysed and concentrated into the reaction buffer for the *in vitro* deSUMOylation assay. This assay was conducted to test the enzymatic activity of the AT60 SUMO protease on the SUMOylated His:JAZ6 substrate. Furthermore, the isopeptidase activity of the AT60^{C115S} mutant protein was also tested to determine whether the active site cysteine residue at position 115 was catalytically significant to the protease activity of AT60. The proteins were dialysed and concentrated into 250µl of SUMO protease buffer and measured. Reaction mixes comprising of the purified His:JAZ6 substrate and AT60 WT or AT60^{C115S} mutant protein in a protein mass ratio of 5:1, respectively, were prepared. The differences in reaction volume were made up with SUMO protease buffer. The reactions were incubated at 28°C for 2 hours and 16 hours (overnight) and analysed by SDS-PAGE and immunoblotting using antibodies raised against anti-AtSUMO1 and anti-His. These reaction conditions were selected as they successfully induced the protease activity for other Arabidopsis SUMO proteases including OTS1 and desi3a-1, in similar *in vitro* deSUMOylation assays performed in preliminary studies (Gwyther, 2018 and Yates, 2018). The blots

from the *in vitro* deSUMOylation assay are displayed in Figure 5.4 and presented with the purified, dialysed and concentrated JAZ6 substrate sample analysed on a separate blot as a negative control for comparative measures.

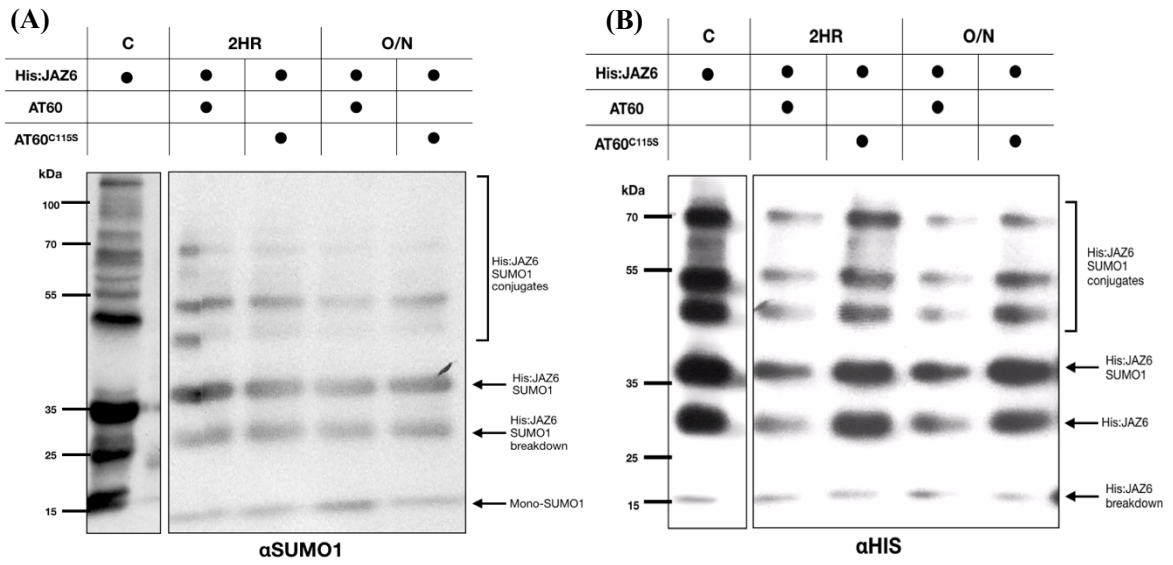


Figure 5.4 – *In vitro* deSUMOylation assay of JAZ6-SUMO conjugation chains by the AT60 DeSi protease. Purified His:JAZ6, AT60 and AT60^{C115S} were dialysed into reaction buffer, concentrated and a reaction mix was prepared containing the His:JAZ6 substrate and AT60 or AT60^{C115S} protein at a ratio of 5:1. All reactions were left to incubate at 28°C for 2 hours (2HR) and 16 hours (overnight = O/N). Reaction samples were then separated by SDS-PAGE gel and transferred for immunoblot analysis probed with anti-AtSUMO1 (A) and anti-His (B). Negative control samples comprising only the purified and concentrated His:JAZ6 protein sample (C) were also analysed on a separate blot for comparative measures. (A) Presents the assay results probed with anti-AtSUMO1 and (B) presents the assay results probed with anti-His. His:JAZ6 = 30.85kDa.

The expected results from this *in vitro* deSUMOylation assay was that in the reaction containing the AT60 WT protease, the JAZ6 SUMO1 conjugation ladder would be significantly more degraded relative to the purified JAZ6 control and AT60^{C115S} mutant reaction, due to the impediment of the protease activity. The results obtained from the *in vitro* deSUMOylation assay were to an extent, consistent with the expected findings. In Figure 5.4-A and Figure 5.4-B, the reactions were analysed by immunoblotting using antibodies raised against AtSUMO1 and His, respectively. In both *in vitro* deSUMOylation results, degradation of the JAZ6 poly-SUMO1 conjugation chain and reduction in the JAZ6 SUMO conjugated protein was observed across all reactions in comparison to the JAZ6 negative control sample. However, the brighter bands detected in the control sample could be attributable to the enhanced staining of the separate gel the samples were run on. It would have been beneficial to have analysed and included the loading controls for the normalisation of the samples.

In Figure 5.4-A, the i*n vitro* deSUMOylation assay containing the AT60 WT protease resulted in a subtle reduction in the JAZ6 SUMO ladder and SUMOylated JAZ6 protein, in comparison to the AT60^{C115S} mutant reaction. In addition, a slight increase in SUMO1 monomers can be observed in the AT60 WT deSUMOylation reaction, which would be expected as the AT60 WT protease would be cleaving and freeing SUMO moieties

from the JAZ6 substrate throughout the deSUMOylation reaction. These findings were more notable in the overnight reactions. Similar results were also observed in Figure 5.4-B. The AT60 WT reaction resulted in a more notable degradation of the JAZ6 poly-SUMO1 conjugation chain and less of the SUMO conjugated JAZ6 protein relative to the $AT60^{C115S}$ mutant reaction. This finding was again more prominent in the overnight deSUMOylation reaction assays. However, more of the His-tagged JAZ6 protein was detected in the mutant deSUMOylation reaction. This finding would have been expected in the AT60 WT reaction as the protease activity of AT60 would be releasing SUMO1 from the JAZ6 substrate increasing the accumulation of His-tagged JAZ6 proteins. However, as the difference in band intensities between the WT and mutant reaction samples and the band intensities within each sample were very subtle, the observed differences could be due to the samples varying in protein concentrations rather than the SUMO protease activity of AT60. For confirmation, the subtle differences in band intensities would need to be quantified and the experiment would need to be repeated. It would have also been beneficial to have analysed and included the loading controls for the normalisation of the samples.

This experiment potentially demonstrated the SUMO protease activity of AT60 in degrading the SUMO-conjugation chain attached to the JAZ6 substrate. Degradation of the JAZ6 poly-SUMO1 conjugates and the SUMO conjugated protein was observed across all reactions relative to the control sample, however this could be due to technical changes rather than biochemical changes. The $AT60^{C115S}$ cysteine active site mutant still exhibited SUMO protease activity, however the efficiency was less than the activity observed in the AT60 WT reaction assays. However, as the band intensities between WT and mutant reactions were very subtle, the differences could be attributable to sample concentration variation rather than AT60 SUMO protease activity. Therefore, although not conclusively, this assay demonstrated the SUMO protease activity of AT60 in cleaving the SUMOylated JAZ6 substrate and provided further evidence in supporting the notion that the cysteine active site within the DeSi motif positioned at the 115 residue specifically in the AT60 SUMO protease, provides the isopeptidase activity required for SUMO deconjugation. Experimental repeats, quantification of band intensity and analysis of loading controls for sample normalisation would be necessary for confirmation.

5.5. Discussion

This chapter explored the biochemical properties of the AT60 DeSi protease through conducting expression and purification analyses, followed by an *in vitro* deSUMOylation assay. In addition, a mutant AT60 protein was generated using site-directed mutagenesis, where the catalytic core cysteine residue at position 115 within the DeSi motif was substituted for the amino acid, serine. This $AT60^{C115S}$ mutant was used to investigate the catalytic function of the DeSi motif. For the *in vitro* deSUMOylation assay, the JAZ6 protein was selected as the target substrate as previous studies had confirmed this protein was commonly SUMOylated in *Arabidopsis* (Srivastava *et al*., 2018). The *E.coli* SS+ cells in which the JAZ6 protein was expressed within, constituted all SUMO conjugation machinery proteins including SUMO1 monomers (Okada *et al.*, 2009). Therefore, this ensured the JAZ6 protein would be heavily SUMOylated for the *in vitro* deSUMOylation assay. Lastly, the deSUMOylation assays were performed to demonstrate the AT60 SUMO protease activity in cleaving SUMO conjugated proteins from the JAZ6 substrate. deSUMOylation reactions consisting of the $AT60^{C115S}$ protease

were also performed to prove the active cysteine residue of the DeSi motif was responsible for the isopeptidase activity of the AT60 DeSi protease.

5.5.1. Protein expression analyses reveal difficulties in optimising expression of AT60 WT and AT60^{C115S} proteins, however, no issues were encountered with expressing the JAZ6 protein

The expression of recombinant AT60 and AT60^{C115S} mutant protein in BL21 (DE3)-RIL cells and JAZ6 in *E. coli* SS+ cells was performed. The expression analysis of the AT60 and AT60^{C115S} protein (Figure 5.2-C-F) displayed the proteins at the expected molecular mass of 29.12kDa. However, the expression analyses revealed that the AT60 proteins often concentrated in inclusion bodies resulting in the protein's absence from the soluble fraction. In *E. coli* cells, inclusion bodies usually form when recombinant proteins are expressed at a high level resulting in protein misfolding, denaturation and hydrophobic aggregation (Thomas and Baneyx, 1996; Fink, 1998). Expression induction conditions were altered as an attempt to slow down protein expression and improve protein folding. Although this resulted in a small proportion of the protein accumulating in the soluble fraction, the majority of the protein was still concentrating in the insoluble fraction. Interestingly, expression analyses of the Arabidopsis SUMO protease, Desi3a, by Orosa *et al*. (2018), found the protease to be stable when expressed, exhibiting no sign of breakdown.

As an attempt to explain the abundance of the AT60 protein in the insoluble fraction, the hydrophobicity and isoelectric point of the protein were considered. Hydropathy plots visualise the hydrophobicity of the protein across the length of the peptide sequence, providing information on the structure of the protein based on its hydrophobicity and hydrophilicity. In the hydropathy plot of AT60, seven peaks were evidently hydrophobic constituting for approximately ~40% of the total peptide sequence. The hydrophobic peaks were situated at the N-terminus and middle section of the protein. With nearly half the AT60 protease scoring positively on the hydropathy index and possessing hydrophobic properties, this offers a biochemical explanation for its abundance in the insoluble fraction. The isoelectric point was also taken into account as a determining factor, as this parameter affects the protein's relative charge, thereby influencing the solubility of the protein. If the environmental pH is equivalent to the protein's isoelectric point, the protein has no net charge and is therefore, at its lowest solubility. The AT60 protease exhibits a small difference between its isoelectric point at 8.12 and the pH level of the *E.coli* cellular environment estimated at around 7.2-7.8 (Wilks and Slonczewski, 2007). Consequently, the AT60 protease will have less charged residues and would be closer to net-neutrality, consequently limiting the protein's solubility, resulting in protein aggregation. Therefore, both the hydrophobicity and isoelectric point of the AT60 protease could be considered contributing factors towards protein aggregation resulting in the minimal accumulation of the protein in the soluble fraction.

The expression analyses of JAZ6 in *E. coli* SS+ cells (Figure 5.2-A-B) displayed the His-tagged JAZ6 protein and SUMO conjugated His-tagged JAZ6 protein at their expected molecular mass of 30.85kDa and 42.85kDa, respectively. The immunoblots were probed with anti-His and anti-SUMO1 antibodies to show definitive SUMOylation of His:JAZ6 expressed in the soluble fraction. The heavier protein bands forming a ladder pattern above the denoted protein bands most likely represent the conjugation of multiple SUMO1 proteins to the His-tagged JAZ6 substrate, producing a poly-SUMO1 chain. For instance, as SUMO1 monomers weigh 11.5kDa, the conjugation of a second SUMO monomer to the SUMOylated His-JAZ6 substrate would result

in an estimated mass of 54.35kDa, which corresponds to the protein band present in the immunoblots at approximately 55kDa. These results were consistent with studies that have conducted *in vitro* SUMOylation assays to generate poly-SUMO chains, which also display a ladder pattern when examined by immunoblot analyses (Castaño-Miquel *et al.*, 2011). More specifically, the findings from this experiment were aligned to results from a paper that re-constituted *in vitro* SUMOylation of JAZ6 tagged with Maltose Binding Protein (MBP) by also exploiting the bacterial SUMO conjugation system; *E.Coli* SS+ cells. The findings from this paper also found a similar ladder pattern representing the poly-SUMO chain conjugated to the JAZ6:MBP substrate (Srivastava *et al.*, 2018). Therefore, the expression of the recombinant JAZ6 protein was optimal for the *in vitro* deSUMOylation assay, where it would act as the target substrate for SUMO deconjugation by the AT60 DeSi protease.

5.5.2. Protein purification analyses successfully purified JAZ6, however issues were encountered when purifying AT60 and AT60^{C115S}

Purification of the AT60, AT60^{C115S} and JAZ6 were performed using the optimal induction conditions established in the protein expression analyses (Table 5.1). The control samples (Figure 5.3-A) and elution fractions (Figure 5.3-C) containing the highest concentration of the respective protein, identified by the UV absorbance levels measured (Figure 5.3-B), were analysed by immunoblotting. With JAZ6, although the SUMOylated His:JAZ6 protein was detected in the binding buffer wash sample, implying a portion of the SUMO conjugated His:JAZ6 protein had not been captured by the HisTrap column, there was strong presence of the protein in the purified elution fractions. The SDS-PAGE gels blotted with anti-His and anti-AtSUMO1, displayed the His:JAZ6 protein at the estimated molecular mass of ~30kDa and SUMOylated His:JAZ6 protein at ~43kDa. The SUMO ladder pattern corresponding to the poly-SUMO1 conjugation chain attached to the JAZ6 substrate could also be observed in both immunoblot results. These findings were again consistent with the paper by Srivastava *et al.* (2018), which detected the SUMOylated His:JAZ6 protein at the estimated molecular mass of ~30kDa in the immunoblot probed with anti-His tag. Therefore, the purified JAZ6 protein would be sufficient enough to act as the substrate for the subsequent *in vitro* deSUMOylation assay.

The purification analyses of the AT60 and AT60^{C115S} proteins were unfortunately not as successful as the purification of JAZ6. Both proteins were identified in the binding buffer wash sample (Figure 5.3-A) indicating the recombinant proteins were not efficiently binding to the HisTrap column. This explained why no protein presence was detected in the purified elution fraction sample (Figure 5.3-C). There were several reasons that could explain the lack of AT60 WT and mutant proteins in the purified elution samples. Firstly, the insolubility of the protein identified during the preliminary expression analyses could be a major determining factor. With a majority of the AT60 protein encapsulated within inclusion bodies, only a small proportion of the soluble protein would be passing through the HisTrap column, reducing the likelihood of the protein being captured by the column during purification. Alternatively, the protein could be unstable in room temperature conditions resulting in protein degradation. For instance, protein purification studies on SUMO protease family members in yeast, ULPs, and in humans, SENPs, were all conducted at 4°C to avoid protein degradation and/or aggregation (Reverter and Lima, 2009). All protein purification analyses were conducted at room temperature, therefore, if more time were available, the purifications for the AT60 proteins would have been performed at 4°C conditions to improve protein stability. In addition, as no protease inhibitor was added to the cell lysis

during protein extraction, as this would impede the protease activity of AT60, endogenous proteolytic and phospholytic enzymes could have degraded the AT60 protease or removed the His tag (Hamilton *et al.*, 2003). Other possible reasons include inadequate binding conditions and obscuration of the His tag due to its location on the AT60 protein (Clark, 1998; Bornhorst and Falke, 2000).

Despite there being no detection of the AT60 WT and AT60^{C115S} protein in the eluted fraction samples, the fractions containing the highest protein concentration determined by the UV absorbance level, were pooled together for dialysis and protein concentration. This was conducted with the aim of increasing the purified protein concentrations of AT60 and AT60^{C115S} protein for the *in vitro* deSUMOylation assay, especially as only a small volume of one eluted fraction was tested for protein presence via immunoblotting analyses.

5.5.3. *In vitro* deSUMOylation assay attempted to demonstrate AT60 SUMO protease activity in degrading the JAZ6 SUMO conjugation chain along with other findings

The *in vitro* deSUMOylation assay was performed to test the SUMO protease activity of the AT60 DeSi protease on the SUMOylated His:JAZ6 substrate. Reactions containing the AT60^{C115S} protease were conducted to investigate the catalytic function of the DeSi motif. The results from the assays did not conclusively demonstrate the AT60 protease activity in degrading the JAZ6 SUMO conjugation chain. The JAZ6 SUMO ladder did appear more depleted in the AT60 and AT60^{C115S} reaction assays in comparison to the negative control sample. There was a reduction in the number of heavier protein bands of the JAZ6 SUMO conjugation chain and considerably fainter protein bands indicating lower protein abundance in the AT60 and AT60^{C115S} assays relative to the JAZ6 negative control sample. However, this could be due to technical changes rather than the SUMO protease activity of AT60 and therefore, experimental repeats would be necessary for confirmation.

These results align with studies that substantiated the SUMO deconjugation activity of different SUMO proteases including OTS1/OTS2 and ESD4 (Sheldon *et al.*, 1999; Conti *et al.*, 2008). The papers all demonstrated this using a similar approach where total proteins derived from WT plants and the respective SUMO protease KO plants (*ots1 ots2* and *esd4-1*) were immunoblotted with anti-AtSUMO1 antibodies to compare SUMO conjugate protein levels. In both studies, the SUMO conjugates increased dramatically in abundance in the KO mutant plants relative to the WT plants, thereby establishing OTS1/OTS2 and ESD4 as bona fide SUMO proteases. The papers also showed a considerable reduction in the abundance of free SUMO monomers in the respective SUMO protease KO mutants in comparison to WT plants, demonstrating the role of OTS1/OTS2 and ESD4 to recycle SUMO monomers from conjugated proteins. This finding was to an extent, aligned with the results from the *in vitro* deSUMOylation assay in this study. The blot visualised with anti-AtSUMO1 showed slight increase in free SUMO1 monomers in the AT60 WT reaction which would be expected if the AT60 protease was actively deconjugating SUMO proteins from the JAZ6 substrate.

Comparing the *in vitro* deSUMOylation assay with the WT AT60 protein and the mutated AT60^{C115S} protein would ascertain whether the AT60 SUMO protease was actively deconjugating the SUMOylated JAZ6 conjugation chain. Moreover, this comparison would validate the cysteine residue situated at position 115

within the DeSi motif, as catalytically important for the isopeptidase activity of the AT60 protease. The ladder of protein bands representing the JAZ6 SUMO conjugation chain were considerably fainter in the WT AT60 reaction, whilst the mutant AT60^{C115S} assay displayed darker and thicker protein bands indicating greater protein abundance. The results were also more prominent in the overnight reactions rather than the 2-hour reactions. However, the JAZ6 SUMO ladder still showed signs of degradation in the mutant AT60^{C115S} reaction relative to the negative control sample, implying the mutant protein was still deconjugating SUMO from the JAZ6 substrate. Furthermore, the difference in degradation levels between the two assays was not significant enough to be conclusive, therefore this variation could be caused by sample concentration differences rather than AT60 SUMO protease activity. In addition, it would also be expected that in the AT60 WT assay, the abundance of the His-tagged JAZ6 protein would be greater than the mutant. However, the results did not reflect this expectation and the experiment would need to be repeated for confirmation.

The *in vitro* deSUMOylation assay results, to an extent, correspond with published literature. Most of the protocols for *in vitro* deSUMOylation assays left the reaction mixture to incubate overnight at 28°C rather than for shorter time periods, as this produced more efficient results consistent with the findings from this experiment (Sheldon *et al.*, 1999; Conti *et al.*, 2008; Orosa *et al.*, 2018). The study by Orosa *et al.* (2018) also conducted an *in vitro* deSUMOylation assay to confirm Desi3a as a bona fide SUMO protease. The researchers generated a mutant protein version of Desi3a with the catalytic core cysteine mutated to serine and subsequently, incubated Desi3a WT and Desi3a^{C168S} with isopeptide-linked poly-SUMO chains. The results displayed no difference in the level of SUMO conjugated proteins in the mutant Desi3a^{C168S} reaction to the poly-SUMO chain control sample, whereas the Desi3a WT reaction significantly degraded the poly-SUMO chains. The Desi3a reaction assay also produced a considerable amount of free SUMO monomers relative to the control and mutant reaction. Furthermore, this paper confirmed FLS2 as a deSUMOylation target of Desi3a, demonstrating a reduction in higher molecular weight SUMO-conjugated isoforms of FLS2 in the reaction with the WT protein and no reduction in the Desi3a^{C168S} containing reaction. Although only a small increase of free SUMO monomers could be observed in the AT60 WT reactions, a significant reduction in the higher molecular weight SUMO-conjugation chain of JAZ6 could be observed in the AT60 WT reaction and not the AT60^{C115S} reaction, coinciding with the findings from the paper by Orosa *et al.* (2018).

As the findings from this experiment did not conclusively demonstrate the SUMO protease activity of the AT60 protein due to the aforementioned limitations, there were several points that would have been addressed if more time were available to ensure the results were conclusive. Firstly, the results do not establish whether the AT60 or AT60^{C115S} proteins were present within the reaction samples. Therefore, this causes ambiguity in the findings from the experiment questioning whether the results were by chance due to variance in exposure times or if they were attributable to the AT60 SUMO protease activity. AT60 and AT60^{C115S} protein control samples post-dialysis and -concentration should have been run next to the assays to clarify their presence in the assays. Secondly, as previously mentioned, it is believed that SUMO proteases provide substrate specificity and therefore, may only target a few SUMOylated substrate proteins. The only published study investigating deSUMOylation of the JAZ6 protein was by Srivastava *et al.* (2018), which found the OTS1/OTS2 protease responsible for the deconjugation of SUMO from the JAZ6 substrate. Therefore, there was no prior knowledge on whether the AT60 protease would be the specific SUMO protease that could deSUMOylate the JAZ6

substrate. Although it was assumed for this experiment that AT60 would have the capability in deconjugating SUMO from JAZ6, to improve the accuracy of the experiment, a truncated version of the AT60 protein containing only the catalytic domain should have been created. This would remove the area of the protein that confers substrate specificity, thereby only testing the deSUMOylation activity of the AT60 DeSi protease (Colby *et al.*, 2006).

5.5.4. Future work in relation to this study

There are several aspects that would have been altered to optimise this experiment if more time were available. The protein purifications for AT60 and AT60^{C115S} proteins would have been performed at 4°C to enhance protein stability and the experiment would have been repeated several times to obtain sufficient amounts of protein for the subsequent *in vitro* deSUMOylation assay. To improve the assays with the AT60^{C115S} mutation protein and fully abolish the isopeptidase enzymatic activity, rather than substituting the active site cysteine residue to serine, which is chemically very similar to one another, mutating the cysteine to another amino acid such as alanine could have provided more insight to the enzymatic activity of AT60 (Lee *et al.*, 2015). Furthermore, as mentioned above, a truncated version of the AT60 protein comprising only of the catalytic domain would have been produced to overcome the issue of target substrate specificity of the AT60 SUMO protease. Lastly, protein biochemistry assays would have also been performed on the AT80 DeSi protease. As determined in section 4, AT80 could be considered the dominant DeSi protease of the DeSi2 sub-group with higher protein expression across all plant organs relative to the AT60 protease. Therefore, the AT80 protein could have been easier to express in the recombinant bacterial system and more optimal for the *in vitro* deSUMOylation assay. Changes to these various parameters would have led to the generation of more conclusive results providing greater insight into the biochemical properties of the AT60 and AT80 DeSi protease.

5.5.5. Final Concluding Remarks

This chapter firstly details how the AT60^{C115S} mutant was successfully produced using site-directed mutagenesis and how all three recombinant proteins were optimally expressed in their respective bacterial systems. Expressing the AT60 WT and AT60^{C115S} mutant proteins in the soluble fraction proved to be harder than the JAZ6 protein due to the protein's isoelectric point and hydrophobicity features. Purification of the recombinant proteins were therefore difficult to achieve especially as the experiment was only conducted twice. The JAZ6 protein was successfully purified, however, the AT60 and AT60^{C115S} mutant protein was not detected in the eluted protein samples. Nevertheless, the elution fractions containing the highest protein concentrations were taken forward for dialysis and concentration as an attempt to increase protein concentration. Lastly, the *in vitro* deSUMOylation assays were conducted with reactions comprising of the purified JAZ6 protein, and AT60 WT or AT60^{C115S} mutant proteins. Although the experiment demonstrated signs of SUMO protease activity of AT60 in degrading the JAZ6 SUMO-conjugation chain, due to several factors mentioned above and as the AT60 mutant still exhibited SUMO protease activity, the results were not fully conclusive. Therefore, this assay provides further evidence to support the notion that the 115 cysteine active site within the DeSi motif specifically in the AT60 SUMO protease, provides the isopeptidase activity required for SUMO deconjugation.

6. Overall Discussion

The overall objective of this thesis was to further our understanding of the deSUMOylation process and the recently discovered class of Arabidopsis SUMO proteases; the DeSis. This study aimed to characterise and define the functionality of two DeSi SUMO proteases, the AT60 and AT80 protease. This was achieved through conducting bioinformatic, genetic and proteomic analyses using online software platforms and laboratory experiments. Transgenic KO mutants and overexpressing transgenic plants of the DeSi proteases were then generated and genotyped, before being subjected to various phenotyping assays for role characterisation. These experiments tested the response of the AT60 and AT80 DeSi proteases to the phytohormone, ABA, and the pathogen response elicitor, flg22. Lastly, protein analyses were conducted to demonstrate the SUMO protease activity of the AT60 DeSi protease.

The DeSi2 subgroup in Arabidopsis comprises of the AT60 and AT80 protease, implying the two DeSi proteases are closely related and share common ancestral lineages. Bioinformatic and proteomic analyses revealed the two proteases are highly homologous to one another, proposing they originated from a gene duplication event and could be functionally redundant to one another. Genetic analyses were conducted to compare the mRNA expression profiles of AT60 and AT80. Differences were observed in their spatial expression pattern, where AT60 was expressed primarily in cauline leaves, whilst AT80 was more abundant in the stem and roots of the Arabidopsis plant. Preliminary subcellular localisation studies were performed by transiently expressing YFP-tagged AT60 and AT80 recombinant proteins in *N. benthamiana* leaves. This study provided some indication that both proteases primarily localised outside the nucleus, specifically along the plasma membrane. The biochemical properties of the AT60 protease were also assessed through performing protein analyses, including an *in vitro* deSUMOylation assay, that demonstrated signs of the SUMO protease isopeptidase activity in degrading SUMO chains from the JAZ6 substrate.

The findings from the genetic and subcellular localisation studies were later corroborated by analyses performed on the single AT60 and AT80 overexpressing lines. These gain-of-function mutants were generated and genotyped through a RT-PCR and real-time PCR, as well as immunoblot analyses, confirming the transgenic lines were overexpressing their respective DeSi protease. Interestingly, the real-time PCR revealed the overexpressing mutants were displaying transcriptional feedback, often observed in functionally redundant genes, where the increase in one DeSi protease resulted in the increase in the other DeSi protease. This further substantiated the theory that the two DeSi proteases could be functionally redundant. Confocal microscopy studies were performed on the single AT60 and AT80 homozygous overexpressing transgenics. This study identified the presence of both proteins outside the nucleus towards the cell periphery, most likely the plasma membrane, aligning with the transient expression findings, indicating the target substrate of both proteases could be membrane-bound. Protein levels of AT80 were overall significantly higher across the seedling, especially in the root system relative to AT60, suggesting unequal functional redundancy may exist between the two highly homologous DeSi proteases, tending towards AT80 being more important. However, as the expression of both proteases are driven by a non-native promoter, the findings should not be considered reflective of the endogenous spatial expression pattern and level of the AT60 and AT80 protease.

The CRISPR/Cas9 binary vector system was used to generate homozygous AT60-AT80KO mutants. As the bioinformatic and phylogenetic analyses confirmed the two proteases were highly homologous to one another and could potentially be functionally redundant, both AT60 and AT80 were knocked-out for mutant transgenic analyses. Due to limitations, the sequencing analysis could only infer and not conclusively confirm the T1 transgenic lines as homozygous CRISPR-Cas9 KO mutants. Nevertheless, taking this into consideration, the T1 lines were propagated to T3 generation. Both KO mutants and overexpressing transgenic lines were subjected to root length and fresh weight assays. The AT60-AT80 double KO mutants exhibited no significant differences in root growth and biomass in comparison to the WT. Whereas, both AT60 and AT80 overexpressing transgenic lines had equivalent root elongation rates to WT, although overall growth was significantly slower, indicating the target substrates of AT60 and AT80 could be involved in the growth and development of Arabidopsis plants.

To further characterise the role of the DeSi proteases, stress response assays were conducted on the transgenic plants. Real-time PCR experiments were first performed to measure the AT60 and AT80 transcript levels in seedlings exposed to ABA and flg22. Localisation studies using confocal microscopy were then performed on the overexpressing transgenic lines to determine the spatial expression pattern of AT60 and AT80 in response to ABA and flg22. In ABA-induced conditions, overall AT60 protein abundance significantly decreased across the seedling aligning with the qPCR findings. In the root length assay, the AT60-AT80KO mutants exhibited hypersensitivity to ABA relative to WT, resulting in significant ABA-induced root growth inhibition. Conversely, AT60 overexpressing lines displayed no difference in ABA-mediated growth arrest in comparison to WT. With the findings consistent with the salt stress phenotypes and expression levels observed in OTS1, it could be postulated that the AT60 DeSi protease acts as a negative regulator of ABA signalling. In Arabidopsis plants, in response to ABA, AT60 levels rapidly decrease, leading to the hyper-SUMOylation of the membrane-bound AT60-target substrate, consequently activating ABA signalling to induce the ABA-mediated growth inhibition phenotype. On the other hand, in response to ABA, AT80 protein abundance significantly increased, particularly in the root tip, coinciding with the real-time PCR results substantiating the unequal redundancy theory. In addition, similar to AT60, AT80 overexpressing lines exhibited no sensitivity difference to ABA, whilst the AT60-AT80KO displayed hypersensitivity to ABA. Although it was clear AT80 was involved in ABA signalling, two theories were proposed linking AT80 DeSi protease and ABA in a signalling pathway. The first speculated that AT80 was functioning homogenously to the AtSPF1 protease in positively regulating ABA signalling and could even be targeting similar target substrates such as ABI5 or MYB30, localised most likely in the root system. Alternatively, AT80 could be negatively regulating ABA signalling like the AT60 and AtOTS1 protease, where the concentration of the SUMOylated AT80-target protein mediates the ABA-induced growth inhibition phenotype. Either way, the two DeSi proteases were shown to have a significant role in regulating ABA signalling in Arabidopsis.

In response to flg22, protein abundance of AT60 significantly reduced across the seedling, aligning with the qPCR results. The AT60-AT80KO mutant lines displayed hypersensitivity to flg22, resulting in significant root growth inhibition relative to WT. Whereas, the AT60 overexpressing lines were insensitive to flg22, exhibiting marginally longer root lengths and significantly greater biomass production in comparison to WT. As the findings from this experiment were consistent with the published findings on Desi3a, it could be

concluded that the AT60 DeSi protease negatively regulates plant immune responses. Upon pathogen perception, AT60 protein abundance reduces leading to the accumulation of SUMOylated membrane-bound AT60-target proteins, which consequently activates early cellular immune responses to the pathogen attack. AT80 protein abundance in response to flg22, considerably reduced in the roots aligning with the qPCR findings, whilst slightly increasing in the leaves, specifically, in the guard cell. The KO mutant lines displayed hypersensitivity to flg22, however, the AT80 overexpressing lines also exhibited increased flg22 sensitivity, however, this was most likely as a result of transgene co-suppression. The findings were consistent with the hypothesised model for AT60, where AT80 negatively regulates plant defence signalling most likely in the root system. However, as AT80 abundance increased in the guard cell in response to flg22, the DeSi protease functionally diverges again from AT60. It could be theorised that upon flg22 perception, AT80 accumulates in the guard cell to deSUMOylate a regulatory membrane-bound protein which in turn induces stomatal closure as part of PTI. Similarly, to ABA responses, it can be concluded that both AT60 and AT80 are implicated in regulating plant immune responses.

6.1. Concluding Remarks and Future Work

This study was successful in identifying key characteristics and the possible functions of the AT60 and AT80 DeSi proteases. The two DeSi proteases are highly homologous to one another sharing very similar characteristics in terms of amino acid sequence and protein structure. AT60 and AT80 proteins were also shown to localise outside the nucleus towards the cell periphery in the plasma membrane. Protein analyses attempted to demonstrate the SUMO protease activity of AT60 in degrading SUMO-conjugated chains from the target substrate, JAZ6. The AT80 DeSi protease had significantly higher expression levels across Arabidopsis seedlings and in response to stress-induced conditions relative to AT60. This negated the theory that the DeSi proteases function redundantly and instead suggested unequal functional redundancy exists between the two proteins, tending towards AT80 being more important. The AT60 DeSi protease had relatively lower protein levels across the seedling, however, had higher expression in the leaves than AT80. In response to ABA and flg22, AT60 levels significantly decreased and phenotyping assays determined AT60 may have a significant role in ABA signalling and plant immune response, most likely as a negative regulator. Higher protein levels of AT80 DeSi protease were observed across the seedling, particularly in the roots, relative to AT60. In response to ABA, AT80 expression increased and stress response assays determined AT80 was implicated in ABA signalling, either as a negative or positive regulator. In the roots, AT80 expression decreased in response to flg22, with the findings from the phenotyping assay substantiating AT80 as a negative regulator of immune responses. Whereas, in the leaf tissue, AT80 accumulated in the guard cells proposing AT80 could be implicated in PTI-associated guard cell responses.

Much more work is needed to substantiate the findings from this study and elucidate the theories hypothesised from the results, in order to fully characterise the AT60 and AT80 DeSi proteases. In particular, the generation and analysis of single AT60 and AT80 KO mutants would have clarified the various speculations and questions that had arisen from this study, including the ability to substantiate the functionally redundant or divergent nature of the two DeSi proteases and validate the regulatory role of each protease in response to ABA and flg22. Genetic complementation assays would be required to confirm whether the DeSi proteases were

responsible for the phenotypes observed in the transgenic lines to confirm their function. Biochemistry assays should be performed on the AT80 DeSi protease as AT80 may be more stable than AT60, and therefore more successful in demonstrating the SUMO activity in deconjugating SUMOylated substrates. The levels of SUMO-conjugation should also be measured in the double KO mutants and overexpressing lines to further validate the role of the two DeSi proteases in the SUMOylation pathway. Lastly, mass spectrometry should be used in future work to correctly identify the target substrates for each DeSi protease as this would provide greater insight into the function and localisation of AT60 and AT80 protease, as well as with conducting future experiments.

This study aimed to elucidate the characteristics and functionality of two newly identified DeSi proteases, AT60 and AT80. With substantial evidence supporting the notion that SUMO proteases regulate the specificity of target substrates, there is increasing importance in studying the role SUMO proteases play in the growth and defence response of model organism plants and in the identification of specific SUMO protease target substrates. With greater knowledge and insight of the SUMO signalling pathway which regulates a vast number of plant adaptive responses, we can further our understanding of the molecular mechanisms by which plants respond to external stress. As a result, this research would significantly contribute towards the development of future breeding and genetic engineering strategies that aim to optimise global food production and alleviate growing concerns of unsustainable population growth and future food shortages.

7. References

Alonso, J. M., Stepanova, A. N., Leisse, T. J., Kim, C. J., Chen, H., Shinn, P., Stevenson, D. K., Zimmerman, J., Barajas, P., Cheuk, R., Gadrinab, C., Heller, C., Jeske, A., Koesema, E., Meyers, C. C., Parker, H., Prednis, L., Ansari, Y., Choy, N., Deen, H., Geralt, M., Hazari, N., Hom, E., Karnes, M., Mulholland, C., Ndubaku, R., Schmidt, I., Guzman, P., Aguilar-Henonin, L., Schmid, M., Weigel, D., Carter, D. E., Marchand, T., Risseeuw, E., Brogden, D., Zeko, A., Crosby, W. L., Berry, C. C. and Ecker, J. R. (2003) 'Genome-wide insertional mutagenesis of Arabidopsis thaliana.', *Science (New York, N.Y.)*, 301(5633), pp. 653–7. doi: 10.1126/science.1086391.

Augustine, R. C. and Vierstra, R. D. (2018) 'SUMOylation: re-wiring the plant nucleus during stress and development.', *Current opinion in plant biology*, 45(Pt A), pp. 143–154. doi: 10.1016/j.pbi.2018.06.006.

Bailey, M., Srivastava, A., Conti, L., Nelis, S., Zhang, C., Florance, H., Love, A., Milner, J., Napier, R., Grant, M. and Sadanandom, A. (2016) 'Stability of small ubiquitin-like modifier (SUMO) proteases OVERLY TOLERANT TO SALT1 and -2 modulates salicylic acid signalling and SUMO1/2 conjugation in Arabidopsis thaliana.', *Journal of experimental botany*. Oxford University Press, 67(1), pp. 353–63. doi: 10.1093/jxb/erv468.

Bartel, B. and Citovsky, V. (2012) 'Focus on ubiquitin in plant biology.', *Plant physiology*. American Society of Plant Biologists, 160(1), p. 1. doi: 10.1104/pp.112.203208.

Beck, M., Wyrsch, I., Strutt, J., Wimalasekera, R., Webb, A., Boller, T. and Robatzek, S. (2014) 'Expression patterns of FLAGELLIN SENSING 2 map to bacterial entry sites in plant shoots and roots', *Journal of Experimental Botany*. Narnia, 65(22), pp. 6487–6498. doi: 10.1093/jxb/eru366.

Benlloch, R. and Lois, L. M. (2018) 'Sumoylation in plants: mechanistic insights and its role in drought stress', *Journal of Experimental Botany*. Oxford University Press, 69(19), pp. 4539–4554. doi: 10.1093/jxb/ery233.

Bernier-Villamor, V., Sampson, D. A., Matunis, M. J. and Lima, C. D. (2002) 'Structural basis for E2-mediated SUMO conjugation revealed by a complex between ubiquitin-conjugating enzyme Ubc9 and RanGAP1.', *Cell*. Elsevier, 108(3), pp. 345–56. doi: 10.1016/S0092-8674(02)00630-X.

Bornhorst, J. A. and Falke, J. J. (2000) 'Purification of proteins using polyhistidine affinity tags.', *Methods in enzymology*, 326, pp. 245–54. Available at: http://www.ncbi.nlm.nih.gov/pubmed/11036646 (Accessed: 10 March 2019).

Bortesi, L. and Fischer, R. (2015) 'The CRISPR/Cas9 system for plant genome editing and beyond', *Biotechnology Advances*. Elsevier, pp. 41–52. doi: 10.1016/j.biotechadv.2014.12.006.

Budhiraja, R., Hermkes, R., Müller, S., Schmidt, J., Colby, T., Panigrahi, K., Coupland, G. and Bachmair, A. (2009) 'Substrates related to chromatin and to RNA-dependent processes are modified by Arabidopsis SUMO isoforms that differ in a conserved residue with influence on desumoylation.', *Plant physiology*. American Society of Plant Biologists, 149(3), pp. 1529–40. doi: 10.1104/pp.108.135053.

van den Burg, H. A., Kini, R. K., Schuurink, R. C. and Takken, F. L. W. (2010) 'Arabidopsis small ubiquitin-like modifier paralogs have distinct functions in development and defense.', *The Plant cell*. American Society of Plant Biologists, 22(6), pp. 1998–2016. doi: 10.1105/tpc.109.070961.

Casneuf, T., Van de Peer, Y. and Huber, W. (2007) 'In situ analysis of cross-hybridisation on microarrays and the inference of expression correlation.', *BMC bioinformatics*. BioMed Central, 8, p. 461. doi: 10.1186/1471-2105-8-461.

Castaño-Miquel, L., Mas, A., Teixeira, I., Seguí, J., Perearnau, A., Thampi, B. N., Schapire, A. L., Rodrigo, N., La Verde, G., Manrique, S., Coca, M. and Lois, L. M. (2017) 'SUMOylation Inhibition Mediated by Disruption of SUMO E1-E2 Interactions Confers Plant Susceptibility to Necrotrophic Fungal Pathogens', *Molecular Plant*. Cell Press, 10(5), pp. 709–720. doi: 10.1016/J.MOLP.2017.01.007.

Castaño-Miquel, L., Seguí, J. and Lois, L. M. (2011) 'Distinctive properties of *Arabidopsis* SUMO paralogues support the *in vivo* predominant role of AtSUMO1/2 isoforms', *Biochemical Journal*, 436(3), pp. 581–590. doi: 10.1042/BJ20101446.

Castaño-Miquel, L., Seguí, J., Manrique, S., Teixeira, I., Carretero-Paulet, L., Atencio, F. and Lois, L. M. (2013) 'Diversification of SUMO-Activating Enzyme in Arabidopsis: Implications in SUMO Conjugation', *Molecular Plant*. Cell Press, 6(5), pp. 1646–1660. doi: 10.1093/MP/SST049.

Castro, P. H., Couto, D., Freitas, S., Verde, N., Macho, A. P., Huguet, S., Botella, M. A., Ruiz-Albert, J., Tavares, R. M., Bejarano, E. R. and Azevedo, H. (2016) 'SUMO proteases ULP1c and ULP1d are required for development and osmotic stress responses in Arabidopsis thaliana', *Plant Molecular Biology*. Springer Netherlands, 92(1–2), pp. 143–159. doi: 10.1007/s11103-016-0500-9.

Castro, P. H., Santos, M. Â., Freitas, S., Cana-Quijada, P., Lourenço, T., Rodrigues, M. A. A., Fonseca, F., Ruiz-Albert, J., Azevedo, J. E., Tavares, R. M., Castillo, A. G., Bejarano, E. R. and Azevedo, H. (2018) 'Arabidopsis thaliana SPF1 and SPF2 are nuclear-located ULP2-like SUMO proteases that act downstream of SIZ1 in plant development', *Journal of Experimental Botany*. Oxford University Press, 69(19), pp. 4633–4649. doi: 10.1093/jxb/ery265.

Castro, P. H., Tavares, R. M., Bejarano, E. R. and Azevedo, H. (2012) 'SUMO, a heavyweight player in plant abiotic stress responses', *Cellular and Molecular Life Sciences*. SP Birkhäuser Verlag Basel, 69(19), pp. 3269–3283. doi: 10.1007/s00018-012-1094-2.

Catala, R., Ouyang, J., Abreu, I. A., Hu, Y., Seo, H., Zhang, X. and Chua, N.-H. (2007) 'The Arabidopsis E3 SUMO ligase SIZ1 regulates plant growth and drought responses.', *The Plant cell*. American Society of Plant Biologists, 19(9), pp. 2952–66. doi: 10.1105/tpc.106.049981.

Chen, H.-C., Hsieh-Feng, V., Liao, P.-C., Cheng, W.-H., Liu, L.-Y., Yang, Y.-W., Lai, M.-H. and Chang, M.-C. (2017) 'The function of OsbHLH068 is partially redundant with its homolog, AtbHLH112, in the regulation of the salt stress response but has opposite functions to control flowering in Arabidopsis', *Plant Molecular Biology*. Springer, 94(4–5), pp. 531–548. doi: 10.1007/s11103-017-0624-6.

Cho, S., Yu, S., Park, J., Mao, Y., Zhu, J.-K., Yun, D.-J. and Lee, B. (2017) 'Accession-Dependent CBF Gene Deletion by CRISPR/Cas System in Arabidopsis', *Frontiers in Plant Science*. Frontiers, 8, p. 1910. doi: 10.3389/fpls.2017.01910.

Chuang, S.-F., Su, L.-H., Cho, C.-C., Pan, Y.-J. and Sun, C.-H. (2012) 'Functional Redundancy of Two Pax-Like Proteins in Transcriptional Activation of Cyst Wall Protein Genes in Giardia lamblia', *PLoS ONE*. Edited by I. Blader. Public Library of Science, 7(2), p. e30614. doi: 10.1371/journal.pone.0030614.

Chuberre, C., Plancot, B., Driouich, A., Moore, J. P., Bardor, M., Gügi, B., & Vicré, M. (2018). Plant Immunity Is Compartmentalized and Specialized in Roots. *Frontiers in Plant Science*, *9*. doi: 10.3389/fpls.2018.01692

Clark, E. D. B. (1998) 'Refolding of recombinant proteins Abbreviations DTT dithiotreitol GdmCI guanidinium chloride', *Current Opinion in Biotechnology*, 9, pp. 157–163. Available at: http://wolfson.huji.ac.il/purification/PDF/Literature/BernardezClark1998.pdf (Accessed: 10 March 2019).

Colby, T. (2006) 'SUMO-Conjugating and SUMO-Deconjugating Enzymes from Arabidopsis', *PLANT PHYSIOLOGY*. American Society of Plant Biologists, 142(1), pp. 318–332. doi: 10.1104/pp.106.085415.

Colby, T., Matthäi, A., Boeckelmann, A. and Stuible, H.-P. (2006) 'SUMO-conjugating and SUMO-deconjugating enzymes from Arabidopsis.', *Plant physiology*. American Society of Plant Biologists, 142(1), pp. 318–32. doi: 10.1104/pp.106.085415.

Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P. D., Wu, X., Jiang, W., Marraffini, L. A. and Zhang, F. (2013) 'Multiplex genome engineering using CRISPR/Cas systems.', *Science (New York, N.Y.)*. NIH Public Access, 339(6121), pp. 819–23. doi: 10.1126/science.1231143.

Conti, L., Kioumourtzoglou, D., O'Donnell, E., Dominy, P. and Sadanandom, A. (2009) 'OTS1 and OTS2 SUMO proteases link plant development and survival under salt stress.', *Plant signaling & behavior*. Taylor & Francis, 4(3), pp. 225–7. Available at: http://www.ncbi.nlm.nih.gov/pubmed/19721757 (Accessed: 1 February 2019).

Conti, L., Nelis, S., Zhang, C., Woodcock, A., Swarup, R., Galbiati, M., Tonelli, C., Napier, R., Hedden, P., Bennett, M. and Sadanandom, A. (2014) 'Small Ubiquitin-like Modifier Protein SUMO Enables Plants to Control Growth Independently of the Phytohormone Gibberellin', *Developmental Cell*. Cell Press, 28(1), pp. 102–110. doi: 10.1016/J.DEVCEL.2013.12.004.

Conti, L., Price, G., O'Donnell, E., Schwessinger, B., Dominy, P. and Sadanandom, A. (2008) 'Small ubiquitin-like modifier proteases OVERLY TOLERANT TO SALT1 and -2 regulate salt stress responses in Arabidopsis.', *The Plant cell*. American Society of Plant Biologists, 20(10), pp. 2894–908. doi: 10.1105/tpc.108.058669.

Earley, K. W., Haag, J. R., Pontes, O., Opper, K., Juehne, T., Song, K. and Pikaard, C. S. (2006) 'Gateway-compatible vectors for plant functional genomics and proteomics', *The Plant Journal*. John Wiley & Sons, Ltd (10.1111), 45(4), pp. 616–629. doi: 10.1111/j.1365-313X.2005.02617.x.

Edwards, K., Johnstone, C. and Thompson, C. (1991) 'A simple and rapid method for the preparation of plant genomic DNA for PCR analysis.', *Nucleic acids research*. Oxford University Press, 19(6), p. 1349. Available at: http://www.ncbi.nlm.nih.gov/pubmed/2030957 (Accessed: 4 December 2018).

Elrouby, N. (2015) 'Analysis of Small Ubiquitin-Like Modifier (SUMO) Targets Reflects the Essential Nature of Protein SUMOylation and Provides Insight to Elucidate the Role of SUMO in Plant Development', *Plant Physiology*. American Society of Plant Biologists, 169(2), pp. 1006–1017. doi: 10.1104/pp.15.01014.

Fink, A. L. (1998) 'Protein aggregation: folding aggregates, inclusion bodies and amyloid', *Folding and Design*, 3(1), pp. R9–R23. doi: 10.1016/S1359-0278(98)00002-9.

Friso, G. and van Wijk, K. J. (2015) 'Update: Post-translational protein modifications in plant metabolism', *Plant Physiology*. Geigenberger, 169, p. pp.01378.2015. doi: 10.1104/pp.15.01378.

Garrido, E., Srivastava, A. K. and Sadanandom, A. (2018) 'Exploiting protein modification systems to boost crop productivity: SUMO proteases in focus', *Journal of Experimental Botany*. Oxford University Press, 69(19), pp. 4625–4632. doi: 10.1093/jxb/ery222.

Gill, G. (2004) 'SUMO and ubiquitin in the nucleus: different functions, similar mechanisms?', *Genes & development*. Cold Spring Harbor Laboratory Press, 18(17), pp. 2046–59. doi: 10.1101/gad.1214604.

Grau-Bové, X., Sebé-Pedrós, A. and Ruiz-Trillo, I. (2015) 'The Eukaryotic Ancestor Had a Complex Ubiquitin Signaling System of Archaeal Origin', *Molecular Biology and Evolution*. Oxford University Press, 32(3), pp. 726–739. doi: 10.1093/molbev/msu334.

Gygi, S. P., Rist, B., Gerber, S. A., Turecek, F., Gelb, M. H. and Aebersold, R. (1999) 'Quantitative analysis of complex protein mixtures using isotope-coded affinity tags', *Nature Biotechnology*, 17(10), pp. 994–999. doi: 10.1038/13690.

Hamilton, J. M. U., Simpson, D. J., Hyman, S. C., Ndimba, B. K. and Slabas, A. R. (2003) 'Ara12 subtilisin-like protease from Arabidopsis thaliana: purification, substrate specificity and tissue localization.', *The Biochemical journal*. Portland Press Limited, 370(Pt 1), pp. 57–67. doi: 10.1042/BJ20021125.

Hashiguchi, A. and Komatsu, S. (2016) 'Impact of Post-Translational Modifications of Crop Proteins under Abiotic Stress.', *Proteomes*. Multidisciplinary Digital Publishing Institute (MDPI), 4(4). doi: 10.3390/proteomes4040042.

Hennig, L., Menges, M., Murray, J. A. H. and Gruissem, W. (2003) *Arabidopsis transcript profiling on Affymetrix GeneChip arrays*, *Plant Molecular Biology*. Available at: http://www.arabidopsis.org (Accessed: 14 March 2019).

Hermkes, R., Fu, Y.-F., Nürrenberg, K., Budhiraja, R., Schmelzer, E., Elrouby, N., Dohmen, R. J., Bachmair, A. and Coupland, G. (2011) 'Distinct roles for Arabidopsis SUMO protease ESD4 and its closest homolog ELS1', *Planta*, 233(1), pp. 63–73. doi: 10.1007/s00425-010-1281-z.

Hickey, C. M., Wilson, N. R. and Hochstrasser, M. (2012) 'Function and regulation of SUMO proteases.', *Nature reviews. Molecular cell biology*. NIH Public Access, 13(12), pp. 755–66. doi: 10.1038/nrm3478.

Hotson, A., Chosed, R., Shu, H., Orth, K. and Mudgett, M. B. (2003) 'Xanthomonas type III effector XopD targets SUMO-conjugated proteins in planta', *Molecular Microbiology*. John Wiley & Sons, Ltd (10.1111), 50(2), pp. 377–389. doi: 10.1046/j.1365-2958.2003.03730.x.

Igarashi, D., Tsuda, K., & Katagiri, F. (2012). The peptide growth factor, phytosulfokine, attenuates pattern-triggered immunity. The Plant Journal, 71(2), 194–204. doi: 10.1111/j.1365-313x.2012.04950.x
Ishida, T., Yoshimura, M., Miura, K. and Sugimoto, K. (2012) 'MMS21/HPY2 and SIZ1, two Arabidopsis SUMO E3 ligases, have distinct functions in development.', *PloS one*. Public Library of Science, 7(10), p. e46897. doi: 10.1371/journal.pone.0046897.

Jiang, F. and Doudna, J. A. (2017) 'CRISPR–Cas9 Structures and Mechanisms', *Annual Review of Biophysics*. Annual Reviews , 46(1), pp. 505–529. doi: 10.1146/annurev-biophys-062215-010822.
Jiang, W., Yang, B. and Weeks, D. P. (2014) 'Efficient CRISPR/Cas9-Mediated Gene Editing in Arabidopsis thaliana and Inheritance of Modified Genes in the T2 and T3 Generations', *PLoS ONE*. Edited by F. Börnke. Public Library of Science, 9(6), p. e99225. doi: 10.1371/journal.pone.0099225.

Jin, J. B., Jin, Y. H., Lee, J., Miura, K., Yoo, C. Y., Kim, W.-Y., Van Oosten, M., Hyun, Y., Somers, D. E., Lee, I., Yun, D.-J., Bressan, R. A. and Hasegawa, P. M. (2007) 'The SUMO E3 ligase, AtSIZ1, regulates flowering by controlling a salicylic acid-mediated floral promotion pathway and through affects on FLC chromatin structure', *The Plant Journal*. John Wiley & Sons, Ltd (10.1111), 53(3), pp. 530–540. doi: 10.1111/j.1365-313X.2007.03359.x.

Jones, D. T., Taylor, W. R. and Thornton, J. M. (1992) 'The rapid generation of mutation data matrices from protein sequences', *Bioinformatics*. Narnia, 8(3), pp. 275–282. doi: 10.1093/bioinformatics/8.3.275.

Jürgen Dohmen, R. (2004) 'SUMO protein modification', *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*. Elsevier, 1695(1–3), pp. 113–131. doi: 10.1016/J.BBAMCR.2004.09.021.

Karan, R. and Subudhi, P. K. (2012) 'A stress inducible SUMO conjugating enzyme gene (SaSce9) from a grass halophyte Spartina alterniflora enhances salinity and drought stress tolerance in Arabidopsis', *BMC Plant Biology*. BioMed Central, 12(1), p. 187. doi: 10.1186/1471-2229-12-187.

Khan, M., Rozhon, W., Unterholzner, S. J., Chen, T., Eremina, M., Wurzinger, B., Bachmair, A., Teige, M., Sieberer, T., Isono, E. and Poppenberger, B. (2014) 'Interplay between phosphorylation and SUMOylation events determines CESTA protein fate in brassinosteroid signalling.', *Nature communications*. Europe PMC Funders, 5, p. 4687. doi: 10.1038/ncomms5687.

Kim, J. Y., Jang, I.-C. and Seo, H. S. (2016) 'COP1 Controls Abiotic Stress Responses by Modulating AtSIZ1 Function through Its E3 Ubiquitin Ligase Activity', *Frontiers in Plant Science*. Frontiers, 7, p. 1182. doi: 10.3389/fpls.2016.01182.

Kim, M.-J., Kim, H. J., Pak, J. H., Cho, H. S., Choi, H. K., Jung, H. W., Lee, D. H. and Chung, Y.-S. (2017) 'Overexpression of *AtSZF2* from *Arabidopsis* Showed Enhanced Tolerance to Salt Stress in Soybean', *Plant Breeding and Biotechnology*. Korean Society of Breeding Science, 5(1), pp. 1–15. doi: 10.9787/PBB.2017.5.1.1.

Kim, S.-I., Park, B. S., Kim, D. Y., Yeu, S. Y., Song, S. I., Song, J. T. and Seo, H. S. (2015) 'E3 SUMO ligase AtSIZ1 positively regulates SLY1-mediated GA signalling and plant development.', *The Biochemical journal*. Portland Press Limited, 469(2), pp. 299–314. doi: 10.1042/BJ20141302.

Kolli, N., Mikolajczyk, J., Drag, M., Mukhopadhyay, D., Moffatt, N., Dasso, M., Salvesen, G. and Wilkinson, K. D. (2010) 'Distribution and paralogue specificity of mammalian deSUMOylating enzymes', *Biochemical Journal*, 430(2), pp. 335–344. doi: 10.1042/BJ20100504.

Kong, X., Luo, X., Qu, G.-P., Liu, P. and Jin, J. B. (2017) '*Arabidopsis* SUMO protease ASP1 positively regulates flowering time partially through regulating FLC stability', *Journal of Integrative Plant Biology*. Wiley/Blackwell (10.1111), 59(1), pp. 15–29. doi: 10.1111/jipb.12509.

Konstantinova, T., Parvanova, D., Atanassov, A. and Djilianov, D. (2003) 'Stable Integration of Transgenes in Tobacco', *Biotechnology & Biotechnological Equipment*, 17(2), pp. 6–14. doi: 10.1080/13102818.2003.10817051.

Kozera, B. and Rapacz, M. (2013) 'Reference genes in real-time PCR.', *Journal of applied genetics*. Springer, 54(4), pp. 391–406. doi: 10.1007/s13353-013-0173-x.

Kurepa, J., Walker, J. M., Smalle, J., Gosink, M. M., Davis, S. J., Durham, T. L., Sung, D.-Y. and Vierstra, R. D. (2003) 'The Small Ubiquitin-like Modifier (SUMO) Protein Modification System in *Arabidopsis*', *Journal of Biological Chemistry*, 278(9), pp. 6862–6872. doi: 10.1074/jbc.M209694200.

Kwon, S. J., Choi, E. Y., Choi, Y. J., Ahn, J. H. and Park, O. K. (2006) 'Proteomics studies of post-translational modifications in plants', *Journal of Experimental Botany*, 57(7), pp. 1547–1551. doi: 10.1093/jxb/erj137.

LeClere, S. and Bartel, B. (2001) 'A library of Arabidopsis 35S-cDNA lines for identifying novel mutants', *Plant Molecular Biology*. Kluwer Academic Publishers, 46(6), pp. 695–703. doi: 10.1023/A:1011699722052.

Lee, E. M., Lee, S. S., Tripathi, B. N., Jung, H. S., Cao, G. P., Lee, Y., Singh, S., Hong, S. H., Lee, K. W., Lee, S. Y., Cho, J.-Y. and Chung, B. Y. (2015) 'Site-directed mutagenesis substituting cysteine for serine in 2-Cys

peroxiredoxin (2-Cys Prx A) of *Arabidopsis thaliana* effectively improves its peroxidase and chaperone functions', *Annals of Botany*, 116(4), pp. 713–725. doi: 10.1093/aob/mcv094.

Lewis, M. and Bell, C. E. (2000) 'A closer view of the conformation of the Lac repressor bound to operator', *Nature Structural Biology*. Nature Publishing Group, 7(3), pp. 209–214. doi: 10.1038/73317.

Li, W. and Ye, Y. (2008) 'Polyubiquitin chains: functions, structures, and mechanisms.', *Cellular and molecular life sciences : CMLS*. NIH Public Access, 65(15), pp. 2397–406. doi: 10.1007/s00018-008-8090-6.

Lin, X.-L., Niu, D., Hu, Z.-L., Kim, D. H., Jin, Y. H., Cai, B., Liu, P., Miura, K., Yun, D.-J., Kim, W.-Y., Lin, R. and Jin, J. B. (2016) 'An Arabidopsis SUMO E3 Ligase, SIZ1, Negatively Regulates Photomorphogenesis by Promoting COP1 Activity.', *PLoS genetics*. Public Library of Science, 12(4), p. e1006016. doi: 10.1371/journal.pgen.1006016.

Liu, L., Jiang, Y., Zhang, X., Wang, X., Wang, Y., Han, Y., Coupland, G., Jin, J. B., Searle, I., Fu, Y.-F. and Chen, F. (2017) 'Two SUMO Proteases SUMO PROTEASE RELATED TO FERTILITY1 and 2 Are Required for Fertility in Arabidopsis.', *Plant physiology*. American Society of Plant Biologists, 175(4), pp. 1703–1719. doi: 10.1104/pp.17.00021.

Liu, X. and Hou, X. (2018) 'Antagonistic Regulation of ABA and GA in Metabolism and Signaling Pathways', *Frontiers in Plant Science*. Frontiers, 9, p. 251. doi: 10.3389/fpls.2018.00251.

Liu, X., Wu, S., Xu, J., Sui, C. and Wei, J. (2017) 'Application of CRISPR/Cas9 in plant biology', *Acta Pharmaceutica Sinica B*. Elsevier, 7(3), pp. 292–302. doi: 10.1016/J.APSB.2017.01.002.

Lockhart, J. (2013) 'Surviving the onslaught: salicylic acid regulates Plasmodesmata closure during pathogen attack in Arabidopsis.', *The Plant cell*. American Society of Plant Biologists, 25(6), p. 1911. doi: 10.1105/tpc.113.250610.

Lois, L. M., Lima, C. D. and Chua, N.-H. (2003) 'Small ubiquitin-like modifier modulates abscisic acid signaling in Arabidopsis.', *The Plant cell*, 15(6), pp. 1347–59. Available at: http://www.ncbi.nlm.nih.gov/pubmed/12782728 (Accessed: 16 March 2019).

Luo, Y.-X., Han, Y.-F., Zhao, Q.-Y., Du, J.-L., Dou, K., Li, L., Chen, S. and He, X.-J. (2018) 'Sumoylation of SUVR2 contributes to its role in transcriptional gene silencing', *Science China Life Sciences*, 61(2), pp. 235–243. doi: 10.1007/s11427-017-9146-2.

Matzke, M., Kanno, T., Daxinger, L., Huettel, B. and Matzke, A. J. (2009) 'RNA-mediated chromatin-based silencing in plants', *Current Opinion in Cell Biology*. Elsevier Current Trends, 21(3), pp. 367–376. doi: 10.1016/J.CEB.2009.01.025.

Mazur, M. J. and van den Burg, H. A. (2012) 'Global SUMO Proteome Responses Guide Gene Regulation, mRNA Biogenesis, and Plant Stress Responses', *Frontiers in Plant Science*. Frontiers, 3, p. 215. doi: 10.3389/fpls.2012.00215.

Millet, Y. A., Danna, C. H., Clay, N. K., Songnuan, W., Simon, M. D., Werck-Reichhart, D., & Ausubel, F. M. (2010). Innate Immune Responses Activated in Arabidopsis Roots by Microbe-Associated Molecular Patterns. *The Plant Cell*, *22*(3), 973–990. doi: 10.1105/tpc.109.069658

Mishra, N., Sun, L., Zhu, X., Smith, J., Srivastava, A. P., Yang, X., Pehlivan, N., Esmaeili, N., Luo, H., Shen, G., Jones, D., Auld, D., Burke, J., Payton, P., Zhang, H., Prakash Srivastava, A., Yang, X., Pehlivan, N., Esmaeili, N., Luo, H., Shen, G., Jones, D., Auld, D., Burke, J., Payton, P. and Zhang, H. (2017) 'Overexpression of the rice SUMO E3 ligase gene OsSIZ1 in cotton enhances drought and heat tolerance, and

substantially improves fiber yields in the field under reduced irrigation and rainfed conditions', *Plant and Cell Physiology*. Oxford University Press, 58(4), pp. 735–746. doi: 10.1093/pcp/pcx032.

Miura, K., Jin, J. B. and Hasegawa, P. M. (2007) 'Sumoylation, a post-translational regulatory process in plants', *Current Opinion in Plant Biology*. Elsevier Current Trends, 10(5), pp. 495–502. doi: 10.1016/J.PBI.2007.07.002.

Miura, K., Jin, J. B., Lee, J., Yoo, C. Y., Stirm, V., Miura, T., Ashworth, E. N., Bressan, R. A., Yun, D.-J. and Hasegawa, P. M. (2007) 'SIZ1-Mediated Sumoylation of ICE1 Controls CBF3/DREB1A Expression and Freezing Tolerance in Arabidopsis', *the Plant Cell Online*, 19(4), pp. 1403–1414. doi: 10.1105/tpc.106.048397.

Miura, K., Lee, J., Jin, J. B., Yoo, C. Y., Miura, T. and Hasegawa, P. M. (2009) 'Sumoylation of ABI5 by the Arabidopsis SUMO E3 ligase SIZ1 negatively regulates abscisic acid signaling', *Proceedings of the National Academy of Sciences*, 106(13), pp. 5418–5423. doi: 10.1073/pnas.0811088106.

Miura, K., Lee, J., Miura, T. and Hasegawa, P. M. (2010) 'SIZ1 controls cell growth and plant development in arabidopsis through salicylic acid', *Plant and Cell Physiology*. Oxford University Press, 51(1), pp. 103–113. doi: 10.1093/pcp/pcp171.

Miura, K., Sato, A., Ohta, M. and Furukawa, J. (2011) 'Increased tolerance to salt stress in the phosphate-accumulating Arabidopsis mutants siz1 and pho2', *Planta*, 234(6), pp. 1191–1199. doi: 10.1007/s00425-011-1476-y.

Moon, J., Parry, G. and Estelle, M. (2004) 'The Ubiquitin-Proteasome Pathway and Plant Development', *THE PLANT CELL*, 16(12), pp. 3181–3195. doi: 10.1105/tpc.104.161220.

Mukhopadhyay, D. and Dasso, M. (2007) 'Modification in reverse: the SUMO proteases', *Trends in Biochemical Sciences*, 32(6), pp. 286–295. doi: 10.1016/j.tibs.2007.05.002.

Murtas, G., Reeves, P. H., Fu, Y.-F., Bancroft, I., Dean, C. and Coupland, G. (2003) 'A nuclear protease required for flowering-time regulation in Arabidopsis reduces the abundance of SMALL UBIQUITIN-RELATED MODIFIER conjugates.', *The Plant cell*. American Society of Plant Biologists, 15(10), pp. 2308–19. doi: 10.1105/tpc.015487.

Nakabayashi, K., Okamoto, M., Koshiba, T., Kamiya, Y. and Nambara, E. (2005) 'Genome-wide profiling of stored mRNA in Arabidopsis thaliana seed germination: epigenetic and genetic regulation of transcription in seed', *The Plant Journal*, 41(5), pp. 697–709. doi: 10.1111/j.1365-313X.2005.02337.x.

Nayak, A. and Müller, S. (2014) 'SUMO-specific proteases/isopeptidases: SENPs and beyond.', *Genome biology*. BioMed Central, 15(7), p. 422. doi: 10.1186/s13059-014-0422-2.

Novatchkova, M., Tomanov, K., Hofmann, K., Stuible, H.-P. and Bachmair, A. (2012) 'Update on sumoylation: defining core components of the plant SUMO conjugation system by phylogenetic comparison.', *The New phytologist*. Europe PMC Funders, 195(1), pp. 23–31. Available at: http://www.ncbi.nlm.nih.gov/pubmed/22799003 (Accessed: 10 October 2018).

Novy, R. F., Drott, D., Yaeger, K. W. and Mierendorf, R. C. (2001) 'Overcoming the codon bias of E . coli for enhanced protein expression 1 Enhanced sensitivity of ChIP through the application of Pellet Paint'. Available at: https://www.semanticscholar.org/paper/Overcoming-the-codon-bias-of-E-.-coli-for-enhanced-Novy-Drott/c2f69086728794296d7d664e88735a9aaf6a470f (Accessed: 19 March 2019).

Okada, S., Nagabuchi, M., Takamura, Y., Nakagawa, T., Shinmyozu, K., Nakayama, J. and Tanaka, K. (2009) 'Reconstitution of Arabidopsis thaliana SUMO Pathways in E. coli: Functional Evaluation of SUMO

Machinery Proteins and Mapping of SUMOylation Sites by Mass Spectrometry', *Plant and Cell Physiology*. Oxford University Press, 50(6), pp. 1049–1061. doi: 10.1093/pcp/pcp056.

Olaofe, O. O., Buist, K. A., Deen, N. G., van der Hoef, M. A. and Kuipers, J. A. M. (2013) 'Segregation dynamics in dense polydisperse gas-fluidized beds', *Powder Technology*. Elsevier Current Trends, 246, pp. 695–706. doi: 10.1016/j.pbi.2018.06.006.

Olbricht, G., Sardesai, N., Gelvin, S. B., Craig, B. A., Doerge, R. W., Olbricht, G. ;, Sardesai, N. ;, Gelvin, S. B. ;, Craig, B. A. ; and Olbricht, G. R. (2009) 'STATISTICAL METHODS FOR AFFYMETRIX TILING ARRAY DATA Recommended Citation). "STATISTICAL METHODS FOR AFFYMETRIX TILING ARRAY DATA," Conference on Applied Statistics in Agriculture STATISTICAL METHODS FOR AFFYMETRIX TILING ARRAY DATA'. doi: 10.4148/2475-7772.1080.

Orosa, B., Yates, G., Verma, V., Srivastava, A. K., Srivastava, M., Campanaro, A., De Vega, D., Fernandes, A., Zhang, C., Lee, J., Bennett, M. J. and Sadanandom, A. (2018) 'SUMO conjugation to the pattern recognition receptor FLS2 triggers intracellular signalling in plant innate immunity', *Nature Communications*. Nature Publishing Group, 9(1), p. 5185. doi: 10.1038/s41467-018-07696-8.

Park, B. S. and Seo, H. S. (2007) 'Roles of SUMO in Plants.', *Journal of crop Science and Biotechnology*, 11(1), pp. 1–6. Available at: https://pdfs.semanticscholar.org/f29b/a7c081d454162537696080e57111da7b8f3a.pdf (Accessed: 28 December 2017).

Park, H. J., Kim, W. Y., Park, H. C., Lee, S. Y., Bohnert, H. J. and Yun, D. J. (2011) 'SUMO and SUMOylation in plants', *Molecules and Cells*. Korean Society for Molecular and Cellular Biology, pp. 305–316. doi: 10.1007/s10059-011-0122-7.

Pauwels, L., De Clercq, R., Goossens, J., Iñigo, S., Williams, C., Ron, M., Britt, A. and Goossens, A. (2018) 'A Dual sgRNA Approach for Functional Genomics in Arabidopsis thaliana.', *G3 (Bethesda, Md.)*. Genetics Society of America, 8(8), pp. 2603–2615. doi: 10.1534/g3.118.200046.

Peng, J., Carol, P., Richards, D. E., King, K. E., Cowling, R. J., Murphy, G. P. and Harberd, N. P. (1997) 'The Arabidopsis GAI gene defines a signaling pathway that negatively regulates gibberellin responses.', *Genes & development*. Cold Spring Harbor Laboratory Press, 11(23), pp. 3194–205. Available at: http://www.ncbi.nlm.nih.gov/pubmed/9389651 (Accessed: 18 October 2018).

Pickett, B. and Meeks-Wagner, D. (1995) *REVI E W A RTl CLE Seeing Double: Appreciating Genetic Redundancy*, *The Plant Cell*. Available at: http://www.plantcell.org/content/plantcell/7/9/1347.full.pdf (Accessed: 18 March 2019).

Rajeevkumar, S., Anunanthini, P. and Sathishkumar, R. (2015) 'Epigenetic silencing in transgenic plants.', *Frontiers in plant science*. Frontiers Media SA, 6, p. 693. doi: 10.3389/fpls.2015.00693.

Ramakers, C., Ruijter, J. M., Deprez, R. H. L. and Moorman, A. F. M. (2003) 'Assumption-free analysis of quantitative real-time polymerase chain reaction (PCR) data.', *Neuroscience letters*, 339(1), pp. 62–6. Available at: http://www.ncbi.nlm.nih.gov/pubmed/12618301 (Accessed: 9 December 2018).

Reverter, D. and Lima, C. D. (2009) 'Preparation of SUMO proteases and kinetic analysis using endogenous substrates.', *Methods in molecular biology (Clifton, N.J.)*. NIH Public Access, 497, pp. 225–39. doi: 10.1007/978-1-59745-566-4_15.

Sadanandom, A., Ádám, É., Orosa, B., Viczián, A., Klose, C., Zhang, C., Josse, E.-M., Kozma-Bognár, L. and Nagy, F. (2015) 'SUMOylation of phytochrome-B negatively regulates light-induced signaling in Arabidopsis

thaliana.', *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 112(35), pp. 11108–13. doi: 10.1073/pnas.1415260112.

Saracco, S. A., Miller, M. J., Kurepa, J. and Vierstra, R. D. (2007) 'Genetic Analysis of SUMOylation in Arabidopsis: Conjugation of SUMO1 and SUMO2 to Nuclear Proteins Is Essential', *PLANT PHYSIOLOGY*, 145(1), pp. 119–134. doi: 10.1104/pp.107.102285.

Schmid, M., Davison, T. S., Henz, S. R., Pape, U. J., Demar, M., Vingron, M., Schölkopf, B., Weigel, D. and Lohmann, J. U. (2005) 'A gene expression map of Arabidopsis thaliana development', *Nature Genetics*, 37(5), pp. 501–506. doi: 10.1038/ng1543.

Schulte, A., Lorenzen, I., Böttcher, M. and Plieth, C. (2006) 'A novel fluorescent pH probe for expression in plants', *Plant Methods*. BioMed Central, 2(1), p. 7. doi: 10.1186/1746-4811-2-7.

Schulz, S., Chachami, G., Kozaczkiewicz, L., Winter, U., Stankovic-Valentin, N., Haas, P., Hofmann, K., Urlaub, H., Ovaa, H., Wittbrodt, J., Meulmeester, E. and Melchior, F. (2012) 'Ubiquitin-specific protease-like 1 (USPL1) is a SUMO isopeptidase with essential, non-catalytic functions', *EMBO reports*, 13(10), pp. 930–938. doi: 10.1038/embor.2012.125.

Sharma, B., Joshi, D., Yadav, P. K., Gupta, A. K. and Bhatt, T. K. (2016) 'Role of Ubiquitin-Mediated Degradation System in Plant Biology.', *Frontiers in plant science*. Frontiers Media SA, 7, p. 806. doi: 10.3389/fpls.2016.00806.

Sheldon, C. C., Burn, J. E., Perez, P. P., Metzger, J., Edwards, J. A., Peacock, W. J. and Dennis, E. S. (1999) 'The FLF MADS Box Gene: A Repressor of Flowering in Arabidopsis Regulated by Vernalization and Methylation', *THE PLANT CELL ONLINE*. American Society of Plant Biologists, 11(3), pp. 445–458. doi: 10.1105/tpc.11.3.445.

Shin, E. J., Shin, H. M., Nam, E., Kim, W. S., Kim, J. H., Oh, B. H. and Yun, Y. (2012) 'DeSUMOylating isopeptidase: A second class of SUMO protease', *EMBO Reports*. Nature Publishing Group, 13(4), pp. 339–346. doi: 10.1038/embor.2012.3.

Smalle, J. and Vierstra, R. D. (2004) 'THE UBIQUITIN 26S PROTEASOME PROTEOLYTIC PATHWAY', *Annual Review of Plant Biology*, 55(1), pp. 555–590. doi: 10.1146/annurev.arplant.55.031903.141801.

Smyth, D. R. (1997) 'Gene silencing: Cosuppression at a distance', *Current Biology*. Cell Press, 7(12), pp. R793–R796. doi: 10.1016/S0960-9822(06)00407-6.

Srivastava, A. K., Orosa, B., Singh, P., Cummins, I., Walsh, C., Zhang, C., Grant, M., Roberts, M. R., Anand, G. S., Fitches, E. and Sadanandom, A. (2018) 'SUMO Suppresses the Activity of the Jasmonic Acid Receptor CORONATINE INSENSITIVE 1', *Plant Cell Advance Publication*. doi: 10.1105/tpc.18.00036.

Thomas, J. G. and Baneyx, F. (1996) 'Protein folding in the cytoplasm of Escherichia coli: requirements for the DnaK-DnaJ-GrpE and GroEL-GroES molecular chaperone machines', *Molecular Microbiology*, 21(6), pp. 1185–1196. doi: 10.1046/j.1365-2958.1996.651436.x.

Tomanov, K., Hardtke, C., Budhiraja, R., Hermkes, R., Coupland, G. and Bachmair, A. (2013) 'Small Ubiquitin-Like Modifier Conjugating Enzyme with Active Site Mutation Acts as Dominant Negative Inhibitor of SUMO Conjugation in *Arabidopsis* [F]', *Journal of Integrative Plant Biology*, 55(1), pp. 75–82. doi: 10.1111/jipb.12016.

Tomanov, K., Zeschmann, A., Hermkes, R., Eifler, K., Ziba, I., Grieco, M., Novatchkova, M., Hofmann, K., Hesse, H. and Bachmair, A. (2014) 'Arabidopsis PIAL1 and 2 promote SUMO chain formation as E4-type SUMO ligases and are involved in stress responses and sulfur metabolism.', *The Plant cell*. American Society of Plant Biologists, 26(11), pp. 4547–60. doi: 10.1105/tpc.114.131300.

Verma, V., Croley, F. and Sadanandom, A. (2018) 'Fifty shades of SUMO: its role in immunity and at the fulcrum of the growth–defence balance', *Molecular Plant Pathology*, 19(6). doi: 10.1111/mpp.12625.

Vijay-Kumar, S., Bugg, C. E., Wilkinson, K. D., Vierstra, R. D., Hatfield, P. M. and Cook, W. J. (1987) 'Comparison of the three-dimensional structures of human, yeast, and oat ubiquitin.', *The Journal of biological chemistry*, 262(13), pp. 6396–9. Available at: http://www.ncbi.nlm.nih.gov/pubmed/3032965 (Accessed: 14 October 2018).

Waese, J., Fan, J., Pasha, A., Yu, H., Fucile, G., Shi, R., Cumming, M., Kelley, L. A., Sternberg, M. J., Krishnakumar, V., Ferlanti, E., Miller, J., Town, C., Stuerzlinger, W. and Provart, N. J. (2017) 'ePlant: Visualizing and Exploring Multiple Levels of Data for Hypothesis Generation in Plant Biology', *The Plant Cell*, 29(8), pp. 1806–1821. doi: 10.1105/tpc.17.00073.

Wang, Q., Qu, G.-P., Kong, X., Yan, Y., Li, J. and Jin, J. B. (2018) '*Arabidopsis* small ubiquitin-related modifier protease ASP1 positively regulates abscisic acid signaling during early seedling development', *Journal of Integrative Plant Biology*, 60(10), pp. 924–937. doi: 10.1111/jipb.12669.

Wang, Z.-P., Xing, H.-L., Dong, L., Zhang, H.-Y., Han, C.-Y., Wang, X.-C. and Chen, Q.-J. (2015) 'Egg cell-specific promoter-controlled CRISPR/Cas9 efficiently generates homozygous mutants for multiple target genes in Arabidopsis in a single generation', *Genome Biology*. BioMed Central, 16(1), p. 144. doi: 10.1186/s13059-015-0715-0.

Watkins, J. M., Hechler, P. J. and Muday, G. K. (2014) 'Ethylene-induced flavonol accumulation in guard cells suppresses reactive oxygen species and moderates stomatal aperture.', *Plant physiology*. American Society of Plant Biologists, 164(4), pp. 1707–17. doi: 10.1104/pp.113.233528.

Webster, D. E. and Thomas, M. C. (2012) 'Post-translational modification of plant-made foreign proteins; glycosylation and beyond', *Biotechnology Advances*. Elsevier, 30(2), pp. 410–418. doi: 10.1016/J.BIOTECHADV.2011.07.015.

Wilks, J. C. and Slonczewski, J. L. (2007) 'pH of the Cytoplasm and Periplasm of Escherichia coli: Rapid Measurement by Green Fluorescent Protein Fluorimetry', *Journal of Bacteriology*, 189(15), pp. 5601–5607. doi: 10.1128/JB.00615-07.

Wolter, F., Klemm, J. and Puchta, H. (2018) 'Efficient *in planta* gene targeting in Arabidopsis using egg cell-specific expression of the Cas9 nuclease of *Staphylococcus aureus*', *The Plant Journal*, 94(4), pp. 735–746. doi: 10.1111/tpj.13893.

Xing, H.-L., Dong, L., Wang, Z.-P., Zhang, H.-Y., Han, C.-Y., Liu, B., Wang, X.-C. and Chen, Q.-J. (2014) 'A CRISPR/Cas9 toolkit for multiplex genome editing in plants', *BMC Plant Biology*. BioMed Central, 14(1), p. 327. doi: 10.1186/s12870-014-0327-y.

Yates, G. (2018) *Identification of a New Class of SUMO Proteases in Plants, and Investigation Into The Role of SUMOylation in Pathogen Perception*. Durham University. Available at: http://etheses.dur.ac.uk/12494/ (Accessed: 18 March 2019).

Yates, G., Srivastava, A. K. and Sadanandom, A. (2016) 'SUMO proteases: uncovering the roles of deSUMOylation in plants', *Journal of Experimental Botany*. Oxford University Press, 67(9), pp. 2541–2548. doi: 10.1093/jxb/erw092.

Zhang, J., De-oliveira-Ceciliato, P., Takahashi, Y., Schulze, S., Dubeaux, G., Hauser, F., Azoulay-Shemer, T., Tõldsepp, K., Kollist, H., Rappel, W.-J. and Schroeder, J. I. (2018) 'Insights into the Molecular Mechanisms of CO2-Mediated Regulation of Stomatal Movements', *Current Biology*. Cell Press, 28(23), pp. R1356–R1363. doi: 10.1016/J.CUB.2018.10.015.

Zhang, S., Qi, Y., Liu, M. and Yang, C. (2013) 'SUMO E3 Ligase AtMMS21 Regulates Drought Tolerance in *Arabidopsis thaliana* [F]', *Journal of Integrative Plant Biology*. Blackwell Publishing Asia, 55(1), pp. 83–95. doi: 10.1111/jipb.12024.

Zhang, S., Qi, Y. and Yang, C. (2010) 'Arabidopsis SUMO E3 ligase AtMMS21 regulates root meristem development.', *Plant signaling & behavior*. Taylor & Francis, 5(1), pp. 53–5. Available at: http://www.ncbi.nlm.nih.gov/pubmed/20592809 (Accessed: 20 October 2018).

Zheng, Y., Chen, Z., Ma, L. and Liao, C. (2018) 'The Ubiquitin E3 Ligase RHA2b Promotes Degradation of MYB30 in Abscisic Acid Signaling.', *Plant physiology*. American Society of Plant Biologists, 178(1), pp. 428–440. doi: 10.1104/pp.18.00683.

Zheng, Y., Schumaker, K. S. and Guo, Y. (2012) 'Sumoylation of transcription factor MYB30 by the small ubiquitin-like modifier E3 ligase SIZ1 mediates abscisic acid response in Arabidopsis thaliana.', *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 109(31), pp. 12822–7. doi: 10.1073/pnas.1202630109.

Zou, C., Lehti-Shiu, M. D., Thomashow, M. and Shiu, S.-H. (2009) 'Evolution of Stress-Regulated Gene Expression in Duplicate Genes of Arabidopsis thaliana', *PLoS Genetics*. Edited by G. P. Copenhaver. Public Library of Science, 5(7), p. e1000581. doi: 10.1371/journal.pgen.1000581.

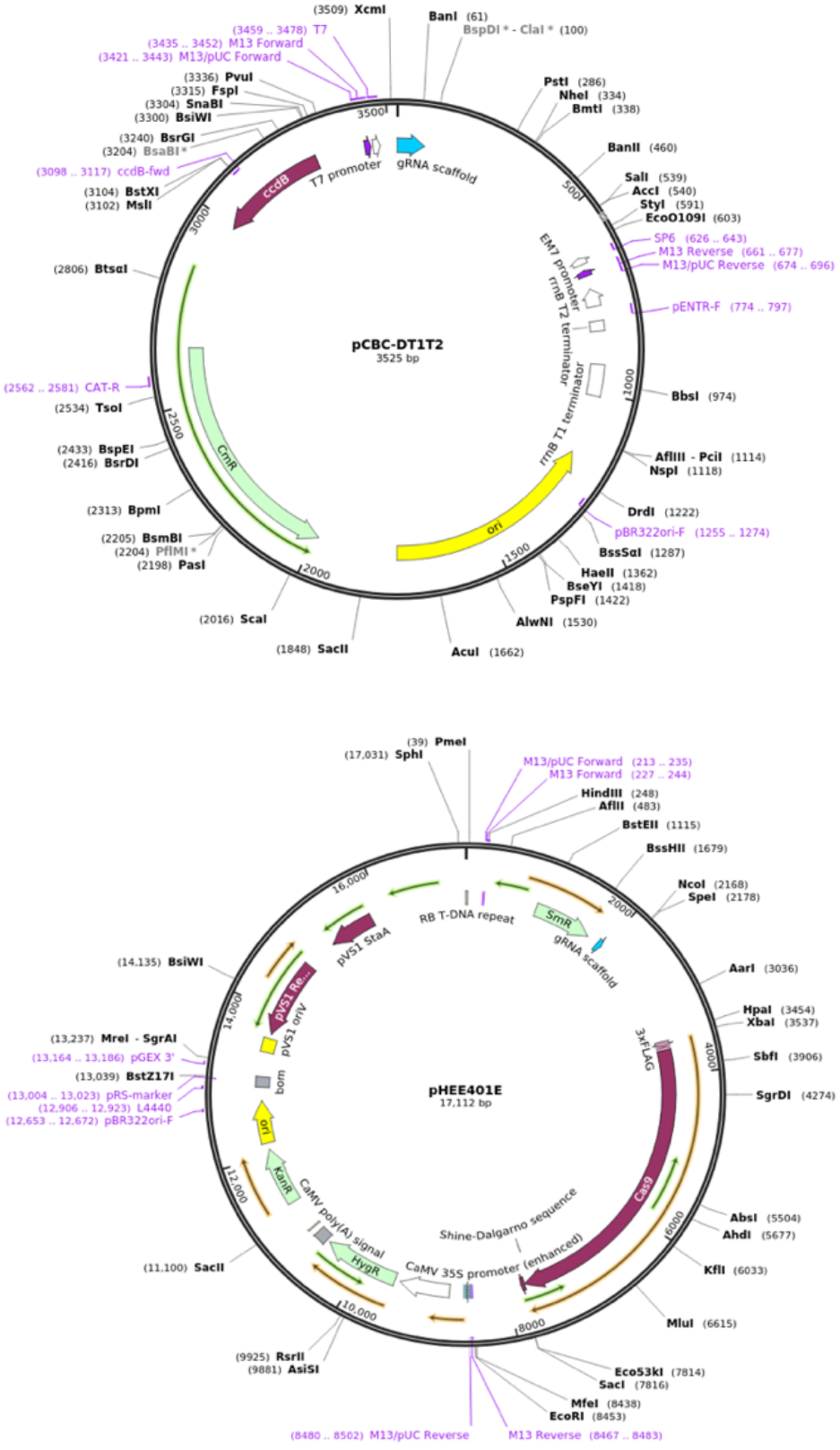
8. Appendix

List of all Primers Used in this Study

Appendix Table 8.1 – List of Primer Sequences

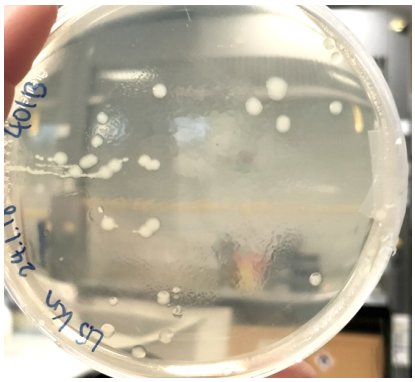
Experimental Use	Primer Name	Sequence (5' to 3')
Assembly of Expression Cassettes	AT60-BsF	ATATATGGTCTCGATTGCGTCTGAGATTGAAAAGTTGTT
	AT60-F0	TGCGTCTGAGATTGAAAAGTTGTTTTAGAGCTAGAAATAGC
	AT60-R0	AACAAAATTGCAATCACTTCTGCAATCTCTTAGTCGACTCTAC
	AT60-BsR	ATTATTGGTCTCGAAACAAAATTGCAATCACTTCTGCAA
	AT80-R0	AACTGAGAAAATCGTCCTTGGCCAATCTCTTAGTCGACTCTAC
	AT80-BsR	ATTATTGGTCTCGAAACTGAGAAAATCGTCCTTGGCCAA
Genotyping AT80	AT80-Fwd	GTACACATACCGTGAGAAAATC
	AT80-Rev	CTGTGTTAAGTAAGCCTTTGAG
Genotyping AT60	AT60-Fwd	GACTAACAACACCATAGTTCAG
	AT80-Rev	TAGAACATCAGAGTCACGGTTC
Genotyping Actin	Actin-Fwd	TCCAAGCTGTTCTCTCCTTGT
	Actin-Rev	CAATCGTGATGACTTGCCCAT
Genotyping YFP coding sequence	YFP-Fwd	ACGTAAACGGCCACAAGTTC
	YFP-Rev	AAG TCG TGC TGC TTC ATG TG
Genotyping T-DNA	SALK T-DNA LB	GCGTGGACCGCTTGCTGCAACT
Genotyping KO Mutants	AT60-G1Fwd	GAGAAAGTGTCGGTTGAGTG
	AT60-G1Rev	CCTCAGTACAGACGAAAAGGG
	AT60-G2Fwd	GGCTCTAATCTAATGGCATATGG
	AT60-G2Rev	GGACTTAACATCCACATGCATTC
	AT80-G2Fwd	GACCTTTGTGCAGGTATATGG
	AT80-G2Rev	GTAAGCCTTTGAGCCAACTA
Real-Time PCR Primers	RT-AT60-Fwd	TGTACTGATCCACCCCAATG
	RT-AT60-Rrev	CACGTTCTGCTGTGAAAAGG
	RT-AT80-Fwd	GCCGTGACTCTGATGTTCTA
	RT-AT80-Rev	GGAGAATAGTGCGTGACTCT
	RT-Actin-Fwd	TGTCCACCTTCCAGCAGATGT
	RT-Actin-Rev	AGCTCAGTAACAGTCCGCCTAG
Site-Directed Mutagenesis	AT60-Ser-Fwd	CTG TCC AAA AAT TCC AAT CAC TTC TGC
	AT60-Ser-Rev	GCAGAAGTGATTGGAATTTTTGGACAG

Plasmid Map of the Vectors Used in the CRISPR/Cas9 Protocol

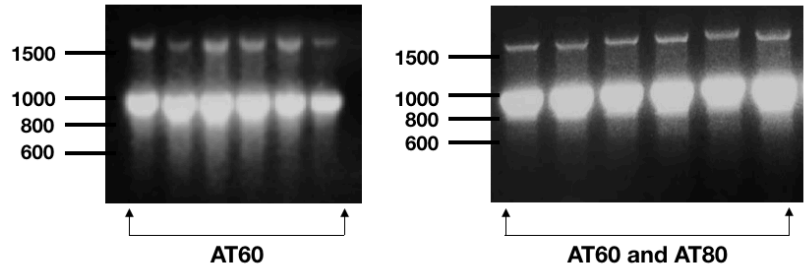


Appendix Figure 8.1 – Plasmid Map of pCBC-DT1T2 and pHEE401E Plasmid

Supplementary Figures from the CRISPR/Cas9 Protocol (Section 3.3 – 3.4)



Appendix Figure 8.2 – Transformed *E. coli* colonies with the sgRNA-containing pHEE401E recombinant plasmid on kanamycin (50µg ml^{-1}) LB agar plates



Appendix Figure 8.3 – Colony PCR products of recombinant *E. coli* cells. M13-Fwd and AT60-BsF primers were used for the colony PCR. The left image displays the colony PCR results for the *E. coli* cells transformed with the pHEE401E plasmid containing the sgRNA targeting the AT60 gene. The right image shows the results from the colony PCR for *E. coli* colonies transformed with the pHEE401E plasmid containing the sgRNA targeting the AT60 and AT80 gene. All colonies were positive as the expected ~1700bp band can be seen for all PCR products indicating successful transformation.

```
pHEE401E      AGCGTGCATAATAAGCCGGTCTCGGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCT      1800
AT60          GATTGCGTCTGAGATTGAAAAGTTGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCT      84
AT60_80       GATTGCGTCTGAGATTGAAAAGTTGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCT      84
                           *          **********************************

pHEE401E      -TGTGAAAATTTGACCATAAGTTTAAATTCTTAAAAAGATATATCTGATCTAGGTGATGG      2394
AT60          -----GAAATTNGACCATAGTTAA---TCCTAAAA----GAATTCTGATTAGGGAN--GN      661
AT60_80       GACTAAAANATTGGCCANGGACGA-TTTTCNCAGT--TTTAGAGCTAGAATAGCAGGTA-      647
                    **  * * ***          * *  *       *   **       *
```

Figure 8.4 – Alignment of the sequenced recombinant plasmids containing spacers sequences of two sgRNAs. Sequencing was performed using the M13-Fwd primer. Sequences for pHEE401E plasmid, AT60 (sgRNA targeting AT60) and AT60_80 (sgRNA dually targeting AT60 and AT80) were aligned by Clustal W. The first yellow box highlights the 1st spacer sequence targeting AT60 which is the same for the AT60 and AT60_80 sample. The second yellow box highlights the 2nd spacer sequence targeting AT60 for the AT60 sample and AT80 for the AT60_80 sample. Both samples contained the correct nucleotide sequence of their respective sgRNA spacer sequence.

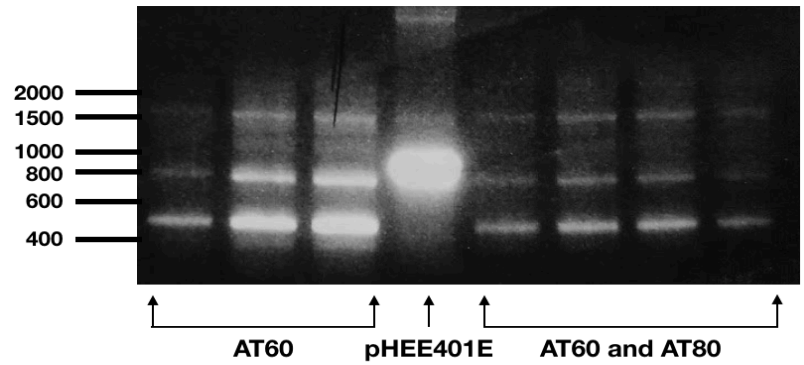


Figure 8.5 – Colony PCR products of recombinant *Agrobacterium* cells. M13-Fwd and AT60-BsF primers were used for the colony PCR. All colonies were positive as the expected approximate 1700bp band can be seen for all PCR products (AT60 and AT60-AT80) indicating successful transformation. As a control, an empty pHEE401E plasmid was run in between the two samples. A faint band can be seen at the expected 1700bp band mark, which is considered contamination.

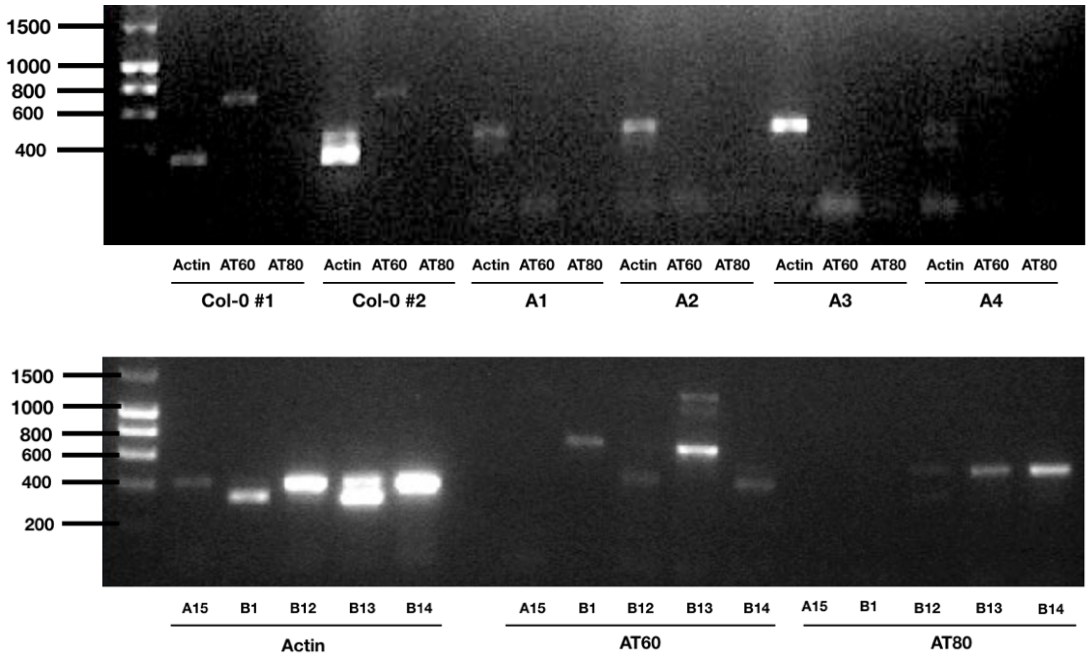


Figure 8.6 – RT-PCR analysis of T1 transgenic KO mutant lines testing the presence of the actin, AT60 and AT80 mRNA transcript. A1-A4 represent the transgenic lines with KOs in the AT60 gene in AT80KO background. B1, B12, B13, and B14 represent the transgenic lines with KOs in the AT60 and AT80 gene in Col-0 background. Col-0 samples were tested as a positive control sample.

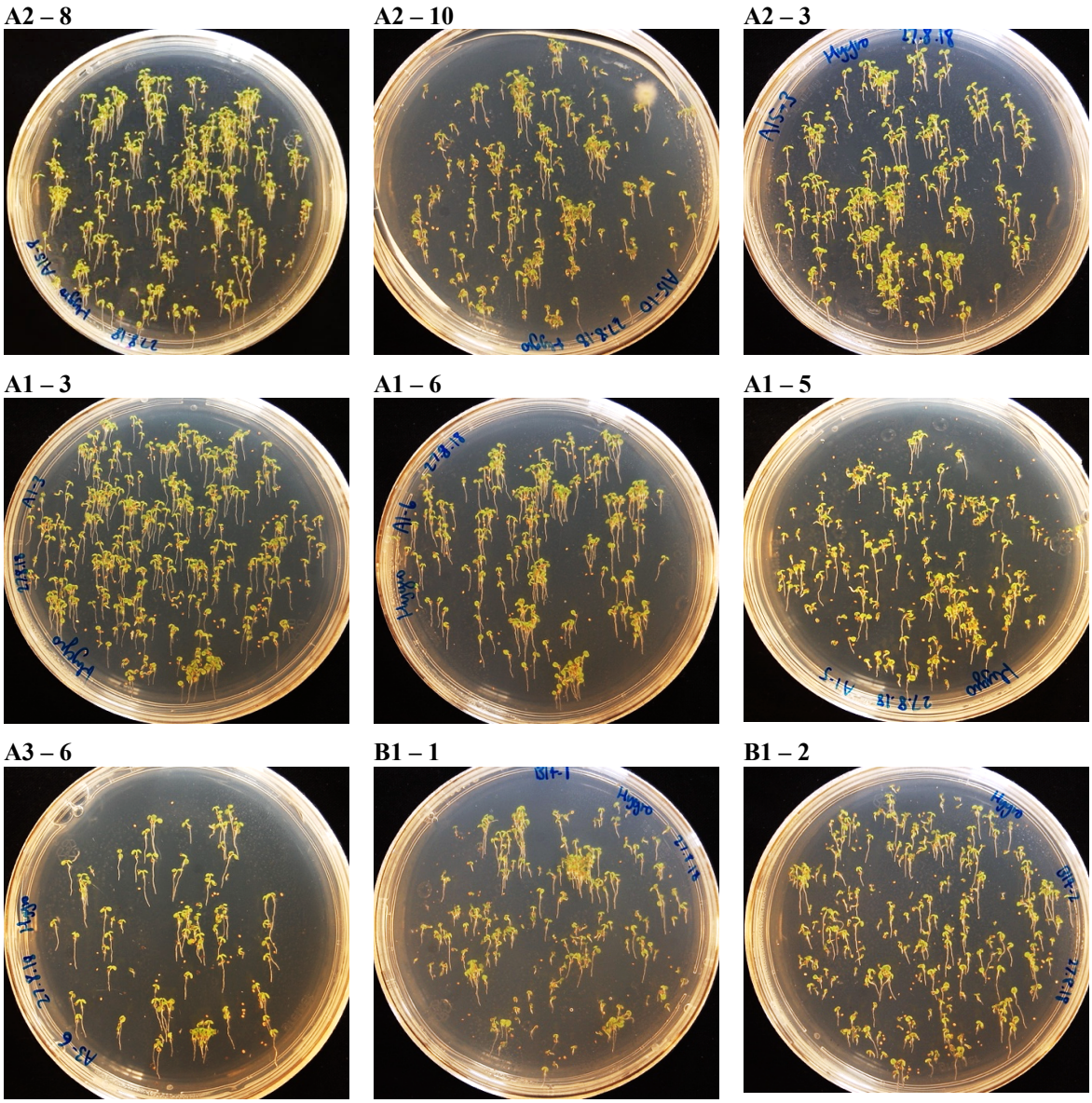


Figure 8.7 – Screening T3 KO transgenic lines for homozygosity on hygromycin (30µg ml^{-1}) MS selection plates

Supplementary Figures from the Confocal Microscopy Study (Section 4.4.4)

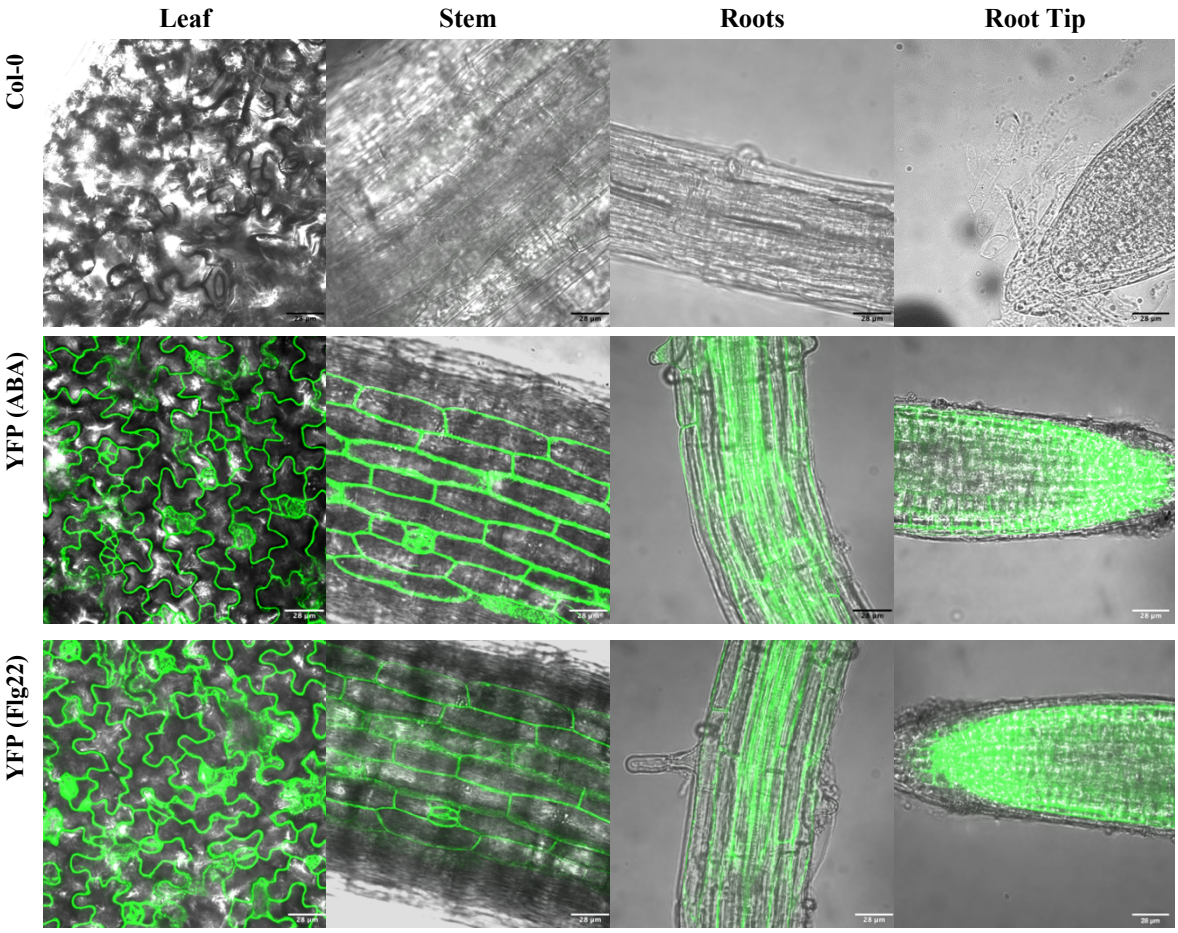


Figure 8.8 – Localisation of the YFP protein in the YFP-only transformed Arabidopsis seedlings in normal and stress-induced growth conditions. The vertical panels in the figure denote the plant tissue of the seedling examined and the horizontal panels indicate the growth conditions. The first displays seedlings grown in normal conditions, followed by seedlings from ABA- (10µM) then flg22- (200nM) induced growth conditions. Scale bar = 28 µm.

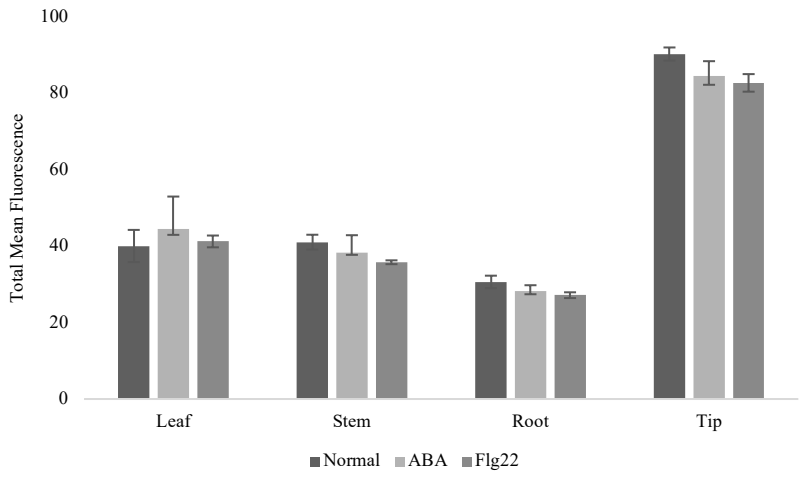


Figure 8.9 – Total mean fluorescence of the YFP protein in the YFP-only transformed Arabidopsis seedlings under normal and stress-induced conditions. Total mean fluorescence of the YFP protein was measured using the Fiji software. Fluorescence of YFP in normal conditions was compared to the fluorescence of the YFP protein in response to ABA (10μM) and flg22 (200nM). Measurements were taken on over 5 images of each section from at least 3 different individual seedlings per genotype for each treatment; $n \geq 14$. Data presented are means ± SE and significance between average fluorescence of the YFP protein in normal and biotic or abiotic stress induced-conditions were assessed using the Mann-Whitney U test. Significance values: * = $P < 0.05$; ** = $P < 0.01$; *** = $P < 0.005$.